Wooley, T. (2019). Nested efficient congruencing and relatives of Vinogradov's mean value theorem. *Proceedings of the London Mathematical Society*, *118*(4), 942-1016. https://doi.org/10.1112/plms.12204

NESTED EFFICIENT CONGRUENCING AND RELATIVES OF VINOGRADOV'S MEAN VALUE THEOREM

TREVOR D. WOOLEY

Abstract. We apply a nested variant of multigrade efficient congruencing to estimate mean values related to that of Vinogradov. We show that when $\varphi_j \in \mathbb{Z}[t]$ $(1 \leqslant j \leqslant k)$ is a system of polynomials with non-vanishing Wronskian, and $s \leqslant k(k+1)/2$, then for all complex sequences $(\mathfrak{a}_n)$, and for each $\varepsilon > 0$, one has

$$\int_{[0,1)^k} \Big| \sum_{|n| \leqslant X} \mathfrak{a}_n e(\alpha_1 \varphi_1(n) + \ldots + \alpha_k \varphi_k(n)) \Big|^{2s} \mathrm{d}\boldsymbol{\alpha} \ll X^{\varepsilon} \Big(\sum_{|n| \leqslant X} |\mathfrak{a}_n|^2 \Big)^s .$$

As a special case of this result, we confirm the main conjecture in Vinogradov's mean value theorem for all exponents k, recovering the recent conclusions of the author (for $k = 3$) and Bourgain, Demeter and Guth (for $k \geqslant 4$). In contrast with the l^2-decoupling method of the latter authors, we make no use of multilinear Kakeya estimates, and thus our methods are of sufficient flexibility to be applicable in algebraic number fields, and in function fields. We outline such extensions.

1. Introduction

This memoir is devoted to a general class of exponential sums and their mean values. We demonstrate how the efficient congruencing method, developed by the author and others starting in late 2010 (see [47]) in the context of Vinogradov systems, may be adapted to handle relatives of Vinogradov's mean value theorem. Indeed, this *nested efficient congruencing method* succeeds in establishing the main conjecture in Vinogradov's mean value theorem for all degrees, a conclusion obtained first by the author in the cubic case (see [55] and arXiv:1401.3150) and subsequently for degrees exceeding three by Bourgain et al. (see [8] and arXiv:1512.01565v3). In contrast with the l^2-decoupling method of the latter authors, nested efficient congruencing makes no use of multilinear Kakeya estimates or other tools apparently intertwined with harmonic analysis in the real setting, and thus our methods are of sufficient flexibility to be applicable in algebraic number fields, and in function fields. We outline such extensions. Further discussion requires that we introduce some notation in order that we pass from descriptive statements to concrete technicalities.

2010 *Mathematics Subject Classification.* 11L15, 11L07, 11P55.

Key words and phrases. Exponential sums, Hardy-Littlewood method, Waring's problem, Strichartz inequalities, congruences.

Given $k \in \mathbb{N}$, we consider polynomials $\varphi_j \in \mathbb{Z}[t]$ $(1 \leqslant j \leqslant k)$ and the associated Wronskian

$$W(t;\boldsymbol{\varphi}) = \det\Big(\varphi_j^{(i)}(t)\Big)_{1\leqslant i,j\leqslant k}. \tag{1.1}$$

Here, following the usual convention, we write $\varphi_j^{(r)}(t)$ for the r-th derivative $\mathrm{d}^r\varphi_j(t)/\mathrm{d}t^r$. A measure of the independence of this system of polynomials $\boldsymbol{\varphi}$ is given by whether or not $W(t;\boldsymbol{\varphi}) = 0$. Our first conclusion supplies an estimate of Strichartz type. As is usual, we write $e(z)$ for $e^{2\pi iz}$.

Theorem 1.1. *Suppose that $\varphi_j \in \mathbb{Z}[t]$ $(1 \leqslant j \leqslant k)$ is a system of polynomials with $W(t;\boldsymbol{\varphi}) \neq 0$. Let s be a positive real number with $s \leqslant k(k+1)/2$. Also, suppose that $(\mathfrak{a}_n)_{n\in\mathbb{Z}}$ is a sequence of complex numbers. Then for each $\varepsilon > 0$, one has*

$$\int_{[0,1)^k}\Big|\sum_{|n|\leqslant X}\mathfrak{a}_n e(\alpha_1\varphi_1(n) + \ldots + \alpha_k\varphi_k(n))\Big|^{2s}\mathrm{d}\boldsymbol{\alpha} \ll X^{\varepsilon}\Big(\sum_{|n|\leqslant X}|\mathfrak{a}_n|^2\Big)^s. \tag{1.2}$$

In particular, under these conditions, one has

$$\int_{[0,1)^k}\Big|\sum_{1\leqslant n\leqslant X} e(\alpha_1\varphi_1(n) + \ldots + \alpha_k\varphi_k(n))\Big|^{2s}\mathrm{d}\boldsymbol{\alpha} \ll X^{s+\varepsilon}. \tag{1.3}$$

We emphasise that here and elsewhere, unless indicated otherwise, the implicit constants in Vinogradov's notation $\ll$ and $\gg$ may depend on ε, s, k and the coefficients of $\boldsymbol{\varphi}$. It follows via orthogonality that when s is a positive integer, then the mean value on the left hand side of (1.3) counts the number of integral solutions of the system of equations

$$\sum_{i=1}^{s}(\varphi_j(x_i) - \varphi_j(y_i)) = 0 \quad (1 \leqslant j \leqslant k), \tag{1.4}$$

with $1 \leqslant x_i, y_i \leqslant X$ $(1 \leqslant i \leqslant s)$. Variants of our methods would yield analogues of Theorem 1.1 in which the polynomials $\varphi_j \in \mathbb{Z}[x]$ are replaced by rational functions lying in $\mathbb{Q}(t)$, or suitably smooth real or p-adic valued functions, with equations replaced by inequalities as appropriate. Likewise, the summands x in (1.3) could be replaced by discretely spaced sets of real or p-adic numbers. We have chosen to provide the most accessible exposition here rather than explore the most general and least transparent framework available to our methods.

By putting $\varphi_j(t) = t^{d_j}$ $(1 \leqslant j \leqslant k)$, we obtain a conclusion on Vinogradov systems in which missing slices are permitted.

Corollary 1.2. *Let $d_1, \ldots, d_k$ be distinct positive integers, and let s be a real number with $0 < s \leqslant k(k+1)/2$. Then for each $\varepsilon > 0$, one has*

$$\int_{[0,1)^k}\Big|\sum_{1\leqslant x\leqslant X} e(\alpha_1 x^{d_1} + \ldots + \alpha_k x^{d_k})\Big|^{2s}\mathrm{d}\boldsymbol{\alpha} \ll X^{s+\varepsilon}. \tag{1.5}$$

Again, when s is a positive integer, it follows via orthogonality that the mean value on the left hand side of (1.5) counts the number of integral solutions of the system

$$\sum_{i=1}^{s}(x_i^{d_j}-y_i^{d_j})=0 \quad (1\leqslant j\leqslant k),$$

with $1\leqslant x_i,y_i\leqslant X$ $(1\leqslant i\leqslant s)$. Aside from recent progress on the Vinogradov system with $d_j=j$ $(1\leqslant j\leqslant k)$, previous published progress on such systems has fallen far short of achieving the range for s delivered by Corollary 1.2. The estimate (1.5) was established for $s\leqslant k+1$ in [44, Theorem 1] via polynomial identities and divisor sum estimates, and indeed such ideas were extended to the setting of Theorem 1.1 for the same range of s in [30, Theorem 1]. We note, however, that extensions to this range have been announced previously. Thus, in 2014 the author announced the proof[1] (via multigrade efficient congruencing) of the upper bound (1.5) in the range $s\leqslant k(k+1)/2-k/3+o(k)$. In addition, Bourgain [6, equation (6.6)] has implicitly announced a result tantamount to the conclusion of Corollary 1.2 in the special case in which $(d_1,\dots,d_k)=(1,2,\dots,k-1,d)$, with $d\geqslant k$. Although one of our purposes in this memoir is to provide a complete published proof for these earlier assertions, we go considerably beyond this earlier work. We remark further that when the degrees d_j are suitably large, the conclusion of Corollary 1.2 follows in a potentially wider range via enhancements of the determinant method of Heath-Brown. Thus, as a consequence of work of the author joint with Salberger (see [34, Theorems 1.3 and 5.2]), one has

$$\int_{[0,1)^k}\Big|\sum_{1\leqslant x\leqslant X}e(\alpha_1x^{d_1}+\ldots+\alpha_kx^{d_k})\Big|^{2s}\mathrm{d}\boldsymbol{\alpha}=s!X^s+O(X^{s-1/2}),$$

provided only that the exponents d_j are distinct and satisfy the condition

$$d_1\cdots d_k\geqslant(2s)^{4s}.$$

The special case of Corollary 1.2 in which $(d_1,\dots,d_k)=(1,2,\dots,k)$ corresponds to the Vinogradov system

$$\sum_{i=1}^{s}(x_i^j-y_i^j)=0 \quad (1\leqslant j\leqslant k). \tag{1.6}$$

Adopting standard notation, we write $J_{s,k}(X)$ for the number of integral solutions of the system (1.6) with $1\leqslant x_i,y_i\leqslant X$ $(1\leqslant i\leqslant s)$. More generally, when s is not necessarily an integer, we put

$$J_{s,k}(X)=\int_{[0,1)^k}\Big|\sum_{1\leqslant x\leqslant X}e(\alpha_1x+\ldots+\alpha_kx^k)\Big|^{2s}\mathrm{d}\boldsymbol{\alpha}. \tag{1.7}$$

[1]See the talk https://www.youtube.com/watch?v=Q5gcLVYqEks from the ELEFANT Workshop, Bonn, July 2014.

The *main conjecture* in Vinogradov's mean value theorem asserts that for each $\varepsilon > 0$, one has

$$J_{s,k}(X) \ll_{\varepsilon,s,k} X^{s+\varepsilon} + X^{2s-k(k+1)/2}. \tag{1.8}$$

Here, we have deviated very slightly from the formulation of the main conjecture asserted in earlier work (see for example [47, equation (1.4)]) by omitting the term ε from the exponent in the second summand on the right hand side of (1.8). This merely recognises the observation, well-known for more than half a century, that the validity of the estimate (1.8) for $s = k(k+1)/2$ implies its validity for all positive real numbers s. Such is evident from an application of Hölder's inequality when $s < k(k+1)/2$, and is immediate from an application of the circle method for $s > k(k+1)/2$. A transparent consequence of Corollary 1.2 yields the main conjecture in full.

Corollary 1.3. *The main conjecture holds in Vinogradov's mean value theorem. Thus, for each $\varepsilon > 0$, one has $J_{s,k}(X) \ll X^{s+\varepsilon} + X^{2s-k(k+1)/2}$. Indeed, when $k \geqslant 3$ and $s > k(k+1)/2$, one has the asymptotic formula*

$$J_{s,k}(X) \sim C_{s,k} X^{2s-k(k+1)/2},$$

where $C_{s,k}$ is a positive number depending at most on s and k.

Theorem 1.1 also delivers the expected Strichartz inequality established for $s \geqslant k(k+1)$ in [59], and subsequently in full in [8].

Corollary 1.4. *Suppose that $k \in \mathbb{N}$, that s is a positive number, and $(\mathfrak{a}_n)_{n\in\mathbb{Z}}$ is a complex sequence. Then, for each positive number ε, one has*

$$\int_{[0,1)^k} \Big| \sum_{|n|\leqslant X} \mathfrak{a}_n e(n\alpha_1 + \ldots + n^k\alpha_k) \Big|^{2s} \mathrm{d}\boldsymbol{\alpha} \ll X^{\varepsilon}(1 + X^{s-k(k+1)/2}) \Big(\sum_{|n|\leqslant X} |\mathfrak{a}_n|^2 \Big)^s.$$

As we have already noted, the main conjecture (1.8) follows for all s from the special case in which $s = k(k+1)/2$. The validity of (1.8) in the case $k = 1$ is of course trivial, and when $k = 2$ the asymptotic formula

$$J_{3,2}(X) \sim \frac{18}{\pi^2} X^3 \log X$$

follows via classical methods (see Rogovskaya [33], and Blomer and Brüdern [5] for sharp versions of this formula). When $k \geqslant 3$ it is expected that the upper bound (1.8) should hold with $\varepsilon = 0$, though at present such is known only when $s \leqslant k+1$ (see [39, Theorem 1]) and $s > k(k+1)/2$ (see Corollary 1.3). The main conjecture (1.8) was established in full for $k = 3$ in the author's previous work [55, Theorem 1.1, Theorem 8.1 and its proof] (see also [18] for an account with certain simplifications). When $k > 3$, the multigrade efficient congruencing method established (1.8) in the range $s \leqslant k(k+1)/2 - k/3 + O(k^{2/3})$, missing the critical exponent $s = k(k+1)/2$ by roughly $k/3$ variables (see [58, Theorem 1.3]). This defect was later remedied in the work of Bourgain et al. [8] by means of their l^2-decoupling method (the reader might refer to [32] for the status of developments at the end of 2016). The nested variant of the multigrade efficient congruencing method that we outline in §2 now also remedies this

defect. Work prior to 2010 preceding the efficient congruencing methods was, meanwhile, far weaker (see for example [23, 41, 43]).

We have already expended considerable space on recording our main conclusions and describing previous results, without pausing to explain the importance of Vinogradov's mean value theorem and its relatives. The recent burst of activity surrounding efficient congruencing, l^2-decoupling, and Vinogradov's mean value theorem offers some justification for this concentration on mean value estimates rather than applications. This is an opportune moment, however, to highlight the central position of Vinogradov's mean value theorem across a large swath of analytic number theory. Current approaches to the asymptotic formula in Waring's problem, the sharpest available estimates for the zero-free region of the Riemann zeta function, and the investigation of equidistribution modulo 1 of polynomial sequences, all depend for their success on estimates associated with Vinogradov's mean value theorem (see [6, 48], [15], [3, 56], respectively). We direct the reader to [50] for an overview of several other applications and an account of recent developments. In §§13 and 14 we record some applications of Theorem 1.1 and its corollaries to Waring's problem and cognate applications.

Several commentators have described the work of Bourgain et al. [8] concerning Vinogradov's mean value theorem as inherently analytic in nature, contrasting it to earlier number-theoretic methods. A comparison of the efficient congruencing methods (see especially [47, 54, 55, 58] and the present paper) with the l^2-decoupling approach [8], however, shows the core of both methods to be strikingly similar. The former applies p-adic short intervals (which is to say, congruence class restrictions) to achieve a p-adic concentration argument via a multiscale iteration, whereas the latter applies real short intervals to achieve the same effect. Number theorists will have little difficulty in translating arguments over $\mathbb{Q}_p$ to analogous arguments over $\mathbb{R}$ (which is to say, over $\mathbb{Q}_\infty$), and vice versa. One of the important messages of the present memoir is that the efficient congruencing ideas are sufficiently flexible that they may be applied to estimate mean values associated with discrete sets of points in a wide variety of fields and their localisations. Examples of such situations include, but are not limited to:

(i) discrete sets of points in $\mathbb{Q}$ and its localisations $\mathbb{Q}_p$ and $\mathbb{R} = \mathbb{Q}_\infty$;
(ii) discrete sets of points in a number field K and its localisations;
(iii) discrete sets of points in $\mathbb{F}_q(t)$ and its localisations;
(iv) discrete sets of points in a function field defined by a curve over a finite field, and its localisations.

In order to illustrate that generalisations of our principal conclusions are easily accessible to our methods, in §15 we establish an analogue of Theorem 1.1 and its corollaries for rings of integers in an arbitrary number field. We also consider function field analogues of our principal conclusions in their most basic form in §17. In both instances, we exploit the relative simplicity of the nested efficient congruencing method as compared to the l^2-decoupling method of Bourgain et al. [8]. We are able, for example, to avoid any discussion of

multilinear Kakeya estimates, the nature of which would be necessarily more mysterious (and presently unknown) in the setting of number fields and function fields. Our approach is also bilinear rather than multilinear, leading to a considerable streamlining of detail. Finally, one inductive aspect of our methods (concerning the number of equations or congruences in play) obviates the need for detailed knowledge of systems of congruences in many variables, as previously supplied by [45]. Thus, the much simpler theory associated with a single congruence in one variable underpins the wider theory without further elaboration. We refer the reader to §§15 and 17 for details, and defer applications to a future paper. We finish by noting that a comprehensive analogue of Theorem 1.1 in the function field setting is the subject of forthcoming work of the author joint with Yu-Ru Liu.

This memoir is organised as follows. We provide a crude outline of the nested efficient congruencing method in §2. Here, readers will find an outline of the ideas new to the nested variant of the multigrade efficient congruencing method, as well as an overview of the efficient congruencing strategy in the large. Sections 3 to 10 inclusive provide a detailed account of the proof of Theorem 3.1, the basic nested inductive step, via the nested efficient congruencing method. The necessary infrastructure is introduced in §3, with a discussion of the implications for translation-dilation invariant families in §4. The nested structure is built inductively, and we introduce the foundation for this induction with the case of a single equation in §5. The general situation is addressed in §§6-9. In §6 we prepare our bilinear structure with some initial conditioning. Then we employ the approximate translation-dilation invariance in §7 so as to generate strong congruence constraints efficiently. In the language of harmonic analysis, this constitutes a *multiscale* aspect to the method. These congruence constraints must be employed iteratively, as described in §8, the analysis of which in §9 prepares the ground for the proof of Theorem 3.1 in §10. Then, in §11, we derive some consequences of Theorem 3.1 for problems involving the number of solutions of congruences in short intervals. In §12 we show how Theorem 3.1 may be employed to deliver Theorem 1.1 and its corollaries. A brief excursion in §13 explores the consequences of Corollary 1.3 for Tarry's problem. An account of the implications of Theorem 1.1 for relatives of Vinogradov's mean value theorem analogous to Hua's lemma is provided in §14, and here we numerically refine some recent work of Bourgain [6] concerning the asymptotic formula in Waring's problem, adding extra details to that exposition. In §15 we indicate how to establish the main conjecture in Vinogradov's mean value theorem for number fields, and we provide some consequences of these conclusions for multivariable analogues of Vinogradov's mean value theorem in §16. Finally, in §17, we confirm the main conjecture in Vinogradov's mean value theorem for function fields in its most basic form.

A couple of notational conventions already deserve mention. Throughout, we make liberal use of vector notation in settings not always conventional in nature. Thus, for example, we write $1 \leqslant \mathbf{x} \leqslant X$ to denote that every coordinate x_i of $\mathbf{x}$ satisfies $1 \leqslant x_i \leqslant X$, and $\mathbf{x} \equiv \xi \pmod{p^h}$ to denote that

$x_i \equiv \xi \pmod{p^h}$ for all indices i. Also, we adopt the convention that when $F : [0,1)^n \to \mathbb{C}$ is integrable, then

$$\oint F(\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha} = \int_{[0,1)^n} F(\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha}. \tag{1.9}$$

Finally, given a situation in which the parameter ε has not already been fixed, in any statement involving the letter ε it is implicitly asserted that the statement holds for each $\varepsilon > 0$. In such circumstances, implicit constants in Vinogradov's notation $\ll$ and $\gg$ may depend on ε.

Acknowledgements: The bulk of this work was completed in February 2016. A period of heavy administration at the University of Bristol slowed the final production of this memoir, but also permitted an evolution of ideas that has greatly simplified several aspects. The author is grateful to the Fields Institute in Toronto for excellent working conditions and support that made the completion of this work possible during the Thematic Program on Unlikely Intersections, Heights, and Efficient Congruencing in the first half of 2017. Further work was supported by the National Science Foundation under Grant No. DMS-1440140 while the author was in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the Spring 2017 semester. The author's work was supported by a European Research Council Advanced Grant under the European Union's Horizon 2020 research and innovation programme via grant agreement No. 695223. Without the release time from the University of Bristol funded by the latter grant, it is difficult to envision that completion of this memoir would have been feasible, and the author wishes to express his gratitude to the ERC for this support.

The author is grateful to Kirsti Biggs and Julia Brandes for numerous suggestions and corrections to an earlier draft of this manuscript. A great debt of gratitude is also owed to the referee for a meticulous and time-consuming review of the paper.

2. A crude outline of nested efficient congruencing

Granted latitude to be economical with technical details and mathematical rigour, we begin in this section by outlining in broad terms the key features of the method central to this paper. The starting point is the translation-dilation invariant (TDI) system of equations (1.6). This TDI property is made evident by the observation that, whenever $\xi \in \mathbb{Z}$ and $q \in \mathbb{N}$, the pair $\mathbf{x}, \mathbf{y}$ satisfies (1.6) if and only if this $2s$-tuple also satisfies the system of equations

$$\sum_{i=1}^{s} \big((qx_i + \xi)^j - (qy_i + \xi)^j\big) = 0 \quad (1 \leqslant j \leqslant k).$$

An application of the binomial theorem rapidly confirms such to be the case.

In the basic version of efficient congruencing (see [47]), one relates the mean value $J_{s,k}(X)$ defined in (1.7) to associated mean values equipped with additional congruence conditions. Write

$$\mathfrak{g}_c(\boldsymbol{\alpha};\xi) = \sum_{\substack{1\leqslant x\leqslant X\\ x\equiv\xi\ (\mathrm{mod}\ p^c)}} e(\alpha_1 x+\ldots+\alpha_k x^k),$$

in which p is a preselected auxiliary prime number of size having order given by a small power of X. Further, for $0\leqslant r\leqslant k$, define the auxiliary mean value

$$K^r_{a,b}(X) = \max_{\xi\not\equiv\eta\ (\mathrm{mod}\ p)} \oint |\mathfrak{g}_a(\boldsymbol{\alpha};\xi)^{2r}\mathfrak{g}_b(\boldsymbol{\alpha};\eta)^{2s-2r}|\,\mathrm{d}\boldsymbol{\alpha}. \tag{2.1}$$

Then it follows via an application of Hölder's inequality that a prime p may be chosen with

$$J_{s,k}(X)\ll p^{2s}K^r_{1,1}(X).$$

It is convenient to possess notation which makes transparent the extent to which the auxiliary mean value $K^r_{a,b}(X)$ exceeds its anticipated magnitude. With this goal in mind, when $s>k(k+1)/2$, we put

$$[\![K^r_{a,b}(X)]\!] = \frac{K^r_{a,b}(X)}{(X/p^a)^{2r-k(k+1)/2}(X/p^b)^{2s-2r}},$$

and when $r\leqslant s\leqslant k(k+1)/2$ we instead define

$$[\![K^r_{a,b}(X)]\!] = \frac{K^r_{a,b}(X)}{(X/p^a)^{r}(X/p^b)^{s-r}}.$$

If, for a given value of s, the mean value $J_{s,k}(X)$ grows approximately like

$$X^{\Lambda}(X^s+X^{2s-k(k+1)/2}),$$

with $\Lambda>0$, then we can infer that for a small number $\varepsilon>0$ one has $[\![K^r_{1,1}(X)]\!]\gg X^{\Lambda-\varepsilon}$. Here, we have made use of the implicit assumption that p has size of order X^θ, where $\theta>0$ is sufficiently small in terms of Λ. Our strategy is now to concentrate this over-abundance of solutions underlying the mean value, relative to the expectation suggested by the main conjecture, through a sequence of auxiliary mean values $K^{r_n}_{a_n,b_n}(X)$ ($n\geqslant 1$). The situation is simplest to describe when $s=k(k+1)$. Here, roughly speaking, one shows that for each $\varepsilon>0$ one has

$$[\![K^k_{a_n,b_n}(X)]\!]\gg X^{\Lambda-\varepsilon}(p^{\psi_n})^{\Lambda}, \tag{2.2}$$

with $a_n\approx k^{n-1}$, $b_n\approx k^n$ and $\psi_n\approx nk^{n-1}(k-1)$. Provided that $\Lambda>0$, then by permitting n to become arbitrarily large sufficiently slowly, one finds that this lower bound for $[\![K^k_{a_n,b_n}(X)]\!]$ vastly exceeds even a trivial estimate for its upper bound, yielding a contradiction. Thus, we are forced to conclude that $\Lambda=0$, and the main conjecture follows for $s\geqslant k(k+1)$.

The lower bound (2.2) is obtained iteratively. By orthogonality, the mean value on the right hand side of (2.1) counts the number of integral solutions

of the system

$$\sum_{i=1}^{r}(x_i^j-y_i^j)=\sum_{l=1}^{s-r}\big((p^bu_l+\eta)^j-(p^bv_l+\eta)^j\big)\quad(1\leqslant j\leqslant k),\tag{2.3}$$

with $1\leqslant \mathbf{x},\mathbf{y}\leqslant X$ and $(1-\eta)/p^b\leqslant \mathbf{u},\mathbf{v}\leqslant (X-\eta)/p^b$, subject to the condition $\mathbf{x}\equiv\mathbf{y}\equiv\xi\pmod{p^a}$. The TDI property of the system (1.6) ensures that (2.3) is equivalent to the system of equations

$$\sum_{i=1}^{r}\big((x_i-\eta)^j-(y_i-\eta)^j\big)=p^{jb}\sum_{l=1}^{s-r}\big(u_l^j-v_l^j\big)\quad(1\leqslant j\leqslant k),$$

and thus one obtains the strong congruence condition

$$\sum_{i=1}^{r}(x_i-\eta)^j\equiv\sum_{i=1}^{r}(y_i-\eta)^j\pmod{p^{jb}}\quad(1\leqslant j\leqslant k).\tag{2.4}$$

One of the technical details suppressed here is the need to condition variables so that $x_1,\dots,x_r$ lie in distinct congruence classes modulo p^{a+1}. This guarantees a level of non-singularity in the solution set that may be exploited via Hensel's lemma.

In the simplest set-up with $s=k(k+1)$ and $r=k$, one shows that for a fixed choice of the k-tuple $\mathbf{y}$ modulo p^{kb}, there are at most $k!(p^{a+b})^{k(k-1)/2}$ possible choices for the k-tuple $\mathbf{x}$ modulo p^{kb}. Since $x_i\equiv\xi\pmod{p^a}$, each variable x_i has at most p^{kb-a} possible choices for its residue class modulo p^{kb}. Reinterpreting the system (2.3) via orthogonality, and applying Hölder's inequality with care, one deduces that

$$\begin{aligned}K_{a,b}^k(X)&\ll(p^{a+b})^{k(k-1)/2}(p^{kb-a})^k\max_{\zeta\not\equiv\eta\pmod p}\oint|\mathfrak{g}_{kb}(\boldsymbol{\alpha};\zeta)^{2k}\mathfrak{g}_b(\boldsymbol{\alpha};\eta)^{2s-2k}|\,\mathrm{d}\boldsymbol{\alpha}\\&\ll p^{\frac12k(k-1)(a+b)+k(kb-a)}K_{b,kb}^k(X)^{\frac{k}{s-k}}J_{s,k}(X/p^b)^{\frac{s-2k}{s-k}}.\end{aligned}\tag{2.5}$$

Here, we have applied the TDI property of the system (1.6) to infer that

$$\oint|\mathfrak{g}_b(\boldsymbol{\alpha};\eta)|^{2s}\,\mathrm{d}\boldsymbol{\alpha}\ll J_{s,k}(X/p^b).$$

The relation (2.5) may be unwound with the estimate

$$J_{s,k}(X/p^b)\ll(X/p^b)^{2s-k(k+1)/2+\Lambda+\varepsilon}.$$

Thus one sees that

$$[\![K_{a,b}^k(X)]\!]\ll X^{\varepsilon}[\![K_{b,kb}^k(X)]\!]^{\frac{k}{s-k}}(X/p^b)^{\frac{s-2k}{s-k}\Lambda},\tag{2.6}$$

and a relation of the shape (2.2) follows.

In the multigrade variant of efficient congruencing, the parameter r in the congruences (2.4) is varied, and we interpret this system in the form

$$\sum_{i=1}^{r}(x_i-\eta)^j\equiv\sum_{i=1}^{r}(y_i-\eta)^j\pmod{p^{(k-r+1)b}}\quad(k-r+1\leqslant j\leqslant k).\tag{2.7}$$

It follows via Hensel's lemma that, when $x_1,\dots,x_r$ lie in distinct congruence classes modulo p^{a+1}, then they are essentially congruent to a permutation of the residue classes $y_1,\dots,y_r$ modulo $p^{(k-r+1)b-(r-1)a}$. A variant of this observation plays a role in the work of the author [16] joint with Ford. Fixing a parameter R with $1 \leqslant R \leqslant k$, this new congruence information may be exploited in a similar manner to that delivering the relation (2.6). One now finds that when $1 \leqslant r \leqslant R$ and $R < s \leqslant k(k+1)/2$, then one has

$$[\![K_{a,b}^r(X)]\!] \ll X^{\varepsilon}[\![K_{a,b}^{r-1}(X)]\!]^{\frac{s-R-r}{s-R-r+1}}[\![K_{b,b'}^R(X)]\!]^{\frac{1}{s-R-r+1}}, \tag{2.8}$$

where $b' = (k-r+1)b-(r-1)a$. Here, we note that when $r=1$, we have $K_{a,b}^{r-1}(X) = K_{a,b}^0(X) \ll J_{s,k}(X/p^b)$, using the TDI property of the system (1.6).

The analysis of this new multigrade efficient congruencing method is necessarily more complicated than with (2.6), since we have numerous quantities $K_{a,b}^r(X)$, with $1 \leqslant r \leqslant R$, in play. The treatment of the iteration involves a complicated tree of possible outcomes. The analysis of [54, 55, 58] iterates the relation (2.8) to obtain a relation of the shape

$$[\![K_{a,b}^R(X)]\!] \ll X^{\varepsilon}(X/p^b)^{\Lambda\phi_0}\prod_{r=1}^{R}[\![K_{b,b_r}^R(X)]\!]^{\phi_r},$$

in which $b_r = (k-r+1)b-(r-1)a$ and the exponents ϕ_r are appropriate positive numbers with $\phi_0+\ldots+\phi_R = 1$. From here one can analyse the tree structure of the iterative process by weighting the outcomes, simplifying to a situation in which one has an averaged relation of the shape

$$[\![K_{a,b}^R(X)]\!] \ll X^{\varepsilon}[\![K_{b,b_r}^R(X)]\!]^{\psi b/b_r}(X/p^b)^{\Lambda\phi_0},$$

for some integer r with $1 \leqslant r \leqslant R$, with ψ a positive number determined by the averaging process, and depending on s and R. Of critical importance is whether or not the exponent ψ exceeds 1. If $\psi > 1$, then the p-adic concentration argument is successful in delivering a contradiction to the assertion that $\Lambda > 0$, much as before, and the main conjecture follows for the value of s in question. On the other hand, if $\psi < 1$, then the iteration fails to deliver a contradiction. It transpires that a choice for R may be made which permits s to be as large as $k(k+1)/2 - k/3 + O(k^{2/3})$, and which allows the main conjecture to be proved for $J_{s,k}(X)$ in this way.

It is at this point that nested efficient congruencing enters the scene. The key observation is that the above iterative processes may be applied without alteration when the system of *equations* (1.6) is replaced by a corresponding system of *congruences*

$$\sum_{i=1}^{s}(x_i^j - y_i^j) \equiv 0 \pmod{p^B} \quad (1 \leqslant j \leqslant k), \tag{2.9}$$

in which B should be interpreted as a large integral parameter. At least, such is the case so long as two subsets of the variables x_i and y_i are restricted to congruence classes modulo p^a and p^b, respectively, in which a and b are sufficiently small that the limitation to a mod p^B environment plays no role in

the above arguments. Such is assured in the system (2.9) when $kb \leqslant B$. Thus, when $s \leqslant k(k+1)/2 - k/3 + O(k^{2/3})$, then in the multigrade argument just described, one may essentially conclude from (2.9) that in an average sense the variables $\mathbf{x}$ and $\mathbf{y}$ are automatically constrained by the additional condition $x_i \equiv y_i \pmod{p^H}$, with $H = \lfloor B/k \rfloor$. More is true. If one has a system of polynomials $\varphi_1, \dots, \varphi_k \in \mathbb{Z}[t]$ with $\varphi_i(t) \equiv t^j \pmod{p^c}$, for some reasonably large parameter c, then the relation (2.9) remains approximately true, since it holds modulo $p^{\min\{c,B\}}$. This may be exploited to show that, in an average sense, one has $x_i \equiv y_i \pmod{p^h}$ with $h = \lfloor c/k \rfloor$. This additional information refines the approximation to the relation (2.9) in a manner similar to the conventional proof of Hensel's lemma. By iterating this idea, one finds even in this more general situation that in an average sense one has $x_i \equiv y_i \pmod{p^H}$.

The observation just sketched may be employed as a substitute for Hensel's lemma in congruence systems of the shape

$$\sum_{i=1}^{u}(x_i - \eta)^j \equiv \sum_{i=1}^{u}(y_i - \eta)^j \pmod{p^{(k-r+1)b}} \quad (k-r+1 \leqslant j \leqslant k),$$

analogous to (2.7), though with u as large as $r(r+1)/2$. The resulting congruence information on the variables x_i and y_i modulo $p^{b'}$, with

$$b' = \lfloor (k-r+1)b/r \rfloor,$$

though weaker than in our earlier treatment, is spread out over many more variables. This, it transpires, offers a sufficient advantage that the earlier defect of $k/3$ variables may be remedied, thereby upgrading the applicability of the multigrade method so as to establish the main conjecture for $J_{s,k}(X)$ when $s \leqslant k(k+1)/2$. The basic approach to the p-adic concentration argument and analysis of the tree of possible outcomes makes use of the same circle of ideas as in the basic multigrade approach. The detailed prosecution of this method will occupy our attention throughout §§3–12.

3. The infrastructure for nested efficient congruencing

We begin by introducing the apparatus required for our proof of Theorem 1.1 via nested efficient congruencing. Although analogous to that of our previous work (see especially [47, 54, 55, 58]) concerning Vinogradov's mean value theorem, we deviate significantly from our previous path. In particular, we incorporate ideas from our work on discrete restriction theory [59] into the method.

Let k be an integer with $k \geqslant 1$, and consider polynomials $\varphi_1, \dots, \varphi_k \in \mathbb{Z}[t]$. Throughout our discussion, we have in mind a fixed prime number p with $p > k$, and a large positive integer B. We require the polynomials φ_j to be sufficiently independent for $1 \leqslant j \leqslant k$. This is achieved by imposing the condition that the system of polynomials $\boldsymbol{\varphi}$ be *p^c-spaced* for an appropriate positive integer c, meaning that

$$\varphi_j(t) \equiv t^j \pmod{p^c} \quad (1 \leqslant j \leqslant k). \tag{3.1}$$

For such a p^c-spaced system of polynomials $\boldsymbol{\varphi}$, one finds that the Wronskian $W(t;\boldsymbol{\varphi})$ defined in (1.1) is necessarily non-zero. Indeed, one has

$$W(t;\boldsymbol{\varphi}) \equiv \det\big(j(j-1)\cdots(j-i+1)t^{j-i}\big)_{1\leqslant i,j\leqslant k} = \prod_{j=1}^{k} j! \not\equiv 0 \pmod{p},$$

whence $W(t;\boldsymbol{\varphi}) \neq 0$. It transpires that the restriction to p^c-spaced systems of polynomials is easily accommodated when establishing such conclusions as Theorem 1.1.

We next define the exponential sums and mean values central to our arguments. Consider a complex sequence $(\mathfrak{a}_n)_{n\in\mathbb{Z}}$ with $\sum_{n\in\mathbb{Z}}|\mathfrak{a}_n| < \infty$. We impose the latter condition for convenience, since this ensures that the Fourier series employed in our arguments are absolutely convergent, and hence that the moments of such series are finite. Formally speaking, all of our arguments apply under the assumption only that $\sum_{n\in\mathbb{Z}}|\mathfrak{a}_n|^r < \infty$ for some $r < 2$, though this requires some interpretation. We have in mind the device of normalising our exponential sums. To this end, when h is a non-negative integer and $\xi \in \mathbb{Z}$, we define $\rho_h(\xi) = \rho_h(\xi;\mathfrak{a})$ by putting

$$\rho_h(\xi;\mathfrak{a}) = \bigg(\sum_{n\equiv\xi \pmod{p^h}} |\mathfrak{a}_n|^2\bigg)^{1/2}. \tag{3.2}$$

Here, we suppress the implicit assumption that the sum is taken over $n \in \mathbb{Z}$. For concision, we write $\rho_0 = \rho_0(\mathfrak{a})$ for $\rho_0(1;\mathfrak{a}) = \big(\sum_{n\in\mathbb{Z}}|\mathfrak{a}_n|^2\big)^{1/2}$. Note that, for all h and ξ, one has $\rho_h(\xi) \leqslant \rho_0 < \infty$.

Next, write

$$\psi(n;\boldsymbol{\alpha}) = \alpha_1\varphi_1(n) + \ldots + \alpha_k\varphi_k(n). \tag{3.3}$$

When h is a non-negative integer and $\xi \in \mathbb{Z}$, we define the exponential sum $\mathfrak{f}_h(\boldsymbol{\alpha};\xi) = \mathfrak{f}_h(\boldsymbol{\alpha};\xi;\mathfrak{a};\boldsymbol{\varphi})$ as follows. When $\rho_h(\xi) > 0$, we put

$$\mathfrak{f}_h(\boldsymbol{\alpha};\xi) = \rho_h(\xi)^{-1} \sum_{n\equiv\xi \pmod{p^h}} \mathfrak{a}_n e(\psi(n;\boldsymbol{\alpha})), \tag{3.4}$$

and otherwise, when $\rho_h(\xi) = 0$, we instead put $\mathfrak{f}_h(\boldsymbol{\alpha};\xi) = 0$. Of course, in the second of these alternatives, one has $\mathfrak{a}_n = 0$ for each $n \in \mathbb{Z}$ with $n \equiv \xi \pmod{p^h}$, and the summation in (3.4) is necessarily 0. It is occasionally useful to abbreviate $\mathfrak{f}_0(\boldsymbol{\alpha};\xi)$ to $f_{\mathfrak{a}}(\boldsymbol{\alpha})$. Note that in the situation in which

$$\mathfrak{a}_n = \begin{cases} 1, & \text{when } 1 \leqslant n \leqslant N, \\ 0, & \text{otherwise,} \end{cases}$$

one has

$$N^{1/2} f_{\mathfrak{a}}(\boldsymbol{\alpha}) = \sum_{1\leqslant n\leqslant N} e(\alpha_1\varphi_1(n) + \ldots + \alpha_k\varphi_k(n)).$$

In order to define the mean values of interest to us, we introduce some concise notation to ease our exposition. We extend the notation (1.9) to accommodate

implicit congruences as follows. Thus, when B is a positive integer, we write

$$\oint_{p^B} F(\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha} = p^{-kB} \sum_{u_1 \bmod p^B} \dots \sum_{u_k \bmod p^B} F(\mathbf{u}/p^B), \tag{3.5}$$

where the summations are taken over complete sets of residues modulo p^B. We then define the mean value $U_{s,k}^B(\mathfrak{a}) = U_{s,k}^{B,\boldsymbol{\varphi}}(\mathfrak{a})$ by putting

$$U_{s,k}^B(\mathfrak{a}) = \oint_{p^B} |f_{\mathfrak{a}}(\boldsymbol{\alpha})|^{2s} d\boldsymbol{\alpha}. \tag{3.6}$$

Note that, by orthogonality, the mean value $U_{s,k}^B(\mathfrak{a})$ counts the integral solutions of the simultaneous congruences

$$\sum_{i=1}^{s} (\varphi_j(x_i) - \varphi_j(y_i)) \equiv 0 \ (\mathrm{mod}\ p^B) \quad (1 \leqslant j \leqslant k), \tag{3.7}$$

with $\mathbf{x}, \mathbf{y} \in \mathbb{Z}$, and with each solution $\mathbf{x}, \mathbf{y}$ being counted with weight

$$\rho_0^{-2s} \prod_{i=1}^{s} \mathfrak{a}_{x_i} \overline{\mathfrak{a}}_{y_i}.$$

We consider also a mean value related to $U_{s,k}^B(\mathfrak{a})$, though with underlying variables restricted to a common congruence class. Thus, we define $U_{s,k}^{B,h}(\mathfrak{a}) = U_{s,k}^{B,h,\boldsymbol{\varphi}}(\mathfrak{a})$ by putting

$$U_{s,k}^{B,h}(\mathfrak{a}) = \rho_0^{-2} \sum_{\xi \bmod p^h} \rho_h(\xi)^2 \oint_{p^B} |\mathfrak{f}_h(\boldsymbol{\alpha};\xi)|^{2s}\,\mathrm{d}\boldsymbol{\alpha}. \tag{3.8}$$

In this instance, it follows via orthogonality that the integral on the right hand side of (3.8) counts the integral solutions $\mathbf{x}, \mathbf{y}$ of the system (3.7) satisfying $\mathbf{x} \equiv \mathbf{y} \equiv \xi \ (\mathrm{mod}\ p^h)$ with weight

$$\rho_h(\xi)^{-2s} \prod_{i=1}^{s} \mathfrak{a}_{x_i} \overline{\mathfrak{a}}_{y_i}.$$

Observe that when $H \geqslant 1$, one may decompose the exponential sum $f_{\mathfrak{a}}(\boldsymbol{\alpha})$ according to residue classes modulo p^H. Thus, it follows from (3.4) that

$$\begin{aligned} f_{\mathfrak{a}}(\boldsymbol{\alpha}) &= \rho_0^{-1} \sum_{\xi \bmod p^H} \sum_{n \equiv \xi \ (\mathrm{mod}\ p^H)} \mathfrak{a}_n e(\psi(n;\boldsymbol{\alpha})) \\ &= \rho_0^{-1} \sum_{\xi \bmod p^H} \rho_H(\xi) \mathfrak{f}_H(\boldsymbol{\alpha};\xi). \end{aligned}$$

An application of Hölder's inequality therefore reveals that

$$|f_{\mathfrak{a}}(\boldsymbol{\alpha})|^{2s} \leqslant \rho_0^{-2s} \Bigg(\sum_{\xi \bmod p^H} 1 \Bigg)^{s} \Bigg(\sum_{\xi \bmod p^H} \rho_H(\xi)^2 \Bigg)^{s-1} \sum_{\xi \bmod p^H} \rho_H(\xi)^2 |\mathfrak{f}_H(\boldsymbol{\alpha};\xi)|^{2s}.$$

Since it follows from (3.2) that

$$\sum_{\xi \bmod p^H} \rho_H(\xi)^2 = \sum_{n\in\mathbb{Z}} |\mathfrak{a}_n|^2 = \rho_0^2,$$

we deduce that

$$|f_{\mathfrak{a}}(\boldsymbol{\alpha})|^{2s} \leqslant \rho_0^{-2} p^{sH} \sum_{\xi \bmod p^H} \rho_H(\xi)^2 |\mathfrak{f}_H(\boldsymbol{\alpha};\xi)|^{2s}. \tag{3.9}$$

We consequently deduce from (3.6) and (3.8) that

$$U_{s,k}^B(\mathfrak{a}) \leqslant p^{sH} U_{s,k}^{B,H}(\mathfrak{a}). \tag{3.10}$$

This relation motivates us to seek an exponent λ having the property that, for each $\varepsilon > 0$, one has

$$U_{s,k}^B(\mathfrak{a}) \ll (p^H)^{\lambda+\varepsilon} U_{s,k}^{B,H}(\mathfrak{a}), \tag{3.11}$$

with as much uniformity in the various parameters as is feasible. It is already apparent from (3.10) that (3.11) holds for some positive number λ satisfying the condition $\lambda \leqslant s$.

In order to make sense of the goal just enunciated, we make some simplifying observations. Observe first that the definition (3.4) of $\mathfrak{f}_h(\boldsymbol{\alpha};\xi)$ is scale invariant with respect to the sequence $(\mathfrak{a}_n)$. Thus, if $\gamma > 0$ and the sequence $(\mathfrak{a}_n)$ is replaced by $(\gamma\mathfrak{a}_n)$, then the exponential sum $\mathfrak{f}_h(\boldsymbol{\alpha};\xi)$ remains unchanged, and likewise therefore the mean value $U_{s,k}^B(\mathfrak{a})$ defined in (3.6) remains unchanged. Denote by $\mathbb{D}$ the set of sequences $(\mathfrak{a}_n)_{n\in\mathbb{Z}}$ with $|\mathfrak{a}_n| \leqslant 1$ $(n \in \mathbb{Z})$ and

$$0 < \sum_{n\in\mathbb{Z}} |\mathfrak{a}_n| < \infty.$$

Also, write $\mathbb{D}_0 = \mathbb{D} \cup \{\mathbf{0}\}$. Then we see that there is no loss of generality in restricting the sequences $(\mathfrak{a}_n)$ under consideration to lie in $\mathbb{D}$.

Next, when $\tau > 0$, denote by $\Phi_\tau(B)$ the set of p^c-spaced k-tuples of polynomials $\boldsymbol{\varphi}$, with $c \geqslant \tau B$. The relation (3.10) ensures that for each fixed $B \in \mathbb{N}$ and $\boldsymbol{\varphi} \in \Phi_\tau(B)$, one has

$$\sup_{(\mathfrak{a}_n)\in\mathbb{D}} \frac{\log(U_{s,k}^B(\mathfrak{a})/U_{s,k}^{B,H}(\mathfrak{a}))}{\log(p^H)} \leqslant s$$

for every non-negative integer H. On the other hand, by considering a sequence $(\mathfrak{b}_n) \in \mathbb{D}$ with $\mathfrak{b}_n = 0$ whenever $n \not\equiv 0 \pmod{p^H}$, one discerns from (3.8) that $U_{s,k}^B(\mathfrak{b}) = U_{s,k}^{B,H}(\mathfrak{b})$, whence

$$\sup_{(\mathfrak{a}_n)\in\mathbb{D}} \frac{\log(U_{s,k}^B(\mathfrak{a})/U_{s,k}^{B,H}(\mathfrak{a}))}{\log(p^H)} \geqslant 0.$$

Given $s > 0$, $\theta \geqslant 1$ and $\tau > 0$, write $H = \lceil B/\theta \rceil$, and define

$$\lambda^*(s,\theta;\tau) = \limsup_{B\to\infty} \sup_{\boldsymbol{\varphi}\in\Phi_\tau(B)} \sup_{(\mathfrak{a}_n)\in\mathbb{D}} \frac{\log(U_{s,k}^B(\mathfrak{a})/U_{s,k}^{B,H}(\mathfrak{a}))}{\log(p^H)}. \tag{3.12}$$

Then we may be assured that $0 \leqslant \lambda^*(s,\theta;\tau) \leqslant s$. Finally, we define the limiting exponent

$$\lambda(s,\theta) = \limsup_{\tau\to 0} \lambda^*(s,\theta;\tau), \tag{3.13}$$

noting that, once again, one has

$$0 \leqslant \lambda(s,\theta) \leqslant s. \tag{3.14}$$

We are now equipped to announce a pivotal estimate that underpins the main inductive step in our argument.

Theorem 3.1. *Suppose that $k \in \mathbb{N}$ and that p is a prime number with $p > k$. Then one has $\lambda(k(k+1)/2, k) = 0$.*

Our methods would show, in fact, that for each natural number k and positive number s, one has

$$\lambda(s,k) = \max\{0, s - k(k+1)/2\}. \tag{3.15}$$

Our main discussion restricts attention to the special case $s = k(k+1)/2$ recorded in Theorem 3.1. This case may be described in a more accessible manner, and is all that is required for the proof of the conclusions recorded in the introduction.

Theorem 3.1 may appear difficult to interpret, and for this reason we present the following corollary.

Corollary 3.2. *Suppose that $k \in \mathbb{N}$ and that p is a prime number with $p > k$. In addition, let $\tau > 0$ and $\varepsilon > 0$. Finally, let B be sufficiently large in terms of k, τ and ε. Put $s = k(k+1)/2$ and $H = \lceil B/k \rceil$. Then for every $\boldsymbol{\varphi} \in \Phi_\tau(B)$, and every sequence $(\mathfrak{a}_n) \in \mathbb{D}_0$, one has*

$$U_{s,k}^{B}(\mathfrak{a}) \ll p^{B\varepsilon} U_{s,k}^{B,H}(\mathfrak{a}). \tag{3.16}$$

We emphasise that the implicit constant in (3.16) may depend on k, τ and ε, but is independent of $\boldsymbol{\varphi}$. We note also that the validity of the relation (3.15) implies that (3.16) holds for all positive numbers $s \leqslant k(k+1)/2$. The conclusion of Corollary 3.2 provides an essentially cost-free concentration towards the diagonal when $s \leqslant k(k+1)/2$. For then the weighted count of solutions of the system (3.7) is dominated by a diagonal contribution in which one may suppose, essentially speaking, that the variables are subject to the additional condition that $\mathbf{x} \equiv \mathbf{y} \pmod{p^{\lceil B/k \rceil}}$. In order to be more concrete concerning this phenomenon, consider the mean value

$$\oint_{p^B} |\mathfrak{f}_H(\boldsymbol{\alpha};\xi)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \tag{3.17}$$

occurring in the definition (3.8) of $U_{s,k}^{B,H}(\mathfrak{a})$. Should this mean value exhibit square-root cancellation, then in view of the normalisation visible in (3.4), it follows from (3.8) that $U_{s,k}^{B,H}(\mathfrak{a}) \ll 1$. The conclusion of Corollary 3.2 then

shows that $U_{s,k}^B(\mathfrak{a}) \ll p^{B\varepsilon}$, which is tantamount to square-root cancellation in the mean value

$$\oint_{p^B} |f_{\mathfrak{a}}(\boldsymbol{\alpha})|^{2s} \, \mathrm{d}\boldsymbol{\alpha}. \tag{3.18}$$

Yet, a priori, the mean value (3.17) is more likely to exhibit square-root cancellation than is (3.18), since the former constrains its underlying variables to an arithmetic progression modulo p^H.

We remark that it would be possible to eliminate the condition $p > k$. Here, two approaches are possible. On the one hand, one could incorporate coefficients divisible by p arising from the extraction of derivatives directly, taking account of the extent to which this inflates subsequent estimates. The impact is modest. Alternatively, one could replace the powers t^j occurring in our definition of p^c-spaced systems by the binomial polynomials $\binom{t+j-1}{j}$.

We next introduce certain auxiliary mean values that play a key role in our arguments. Throughout, we fix $s = k(k+1)/2$. Let a, b, c and ν be non-negative integers. We consider a p^c-spaced k-tuple of polynomials $\boldsymbol{\varphi} \in \mathbb{Z}[t]^k$, and we fix a complex sequence $(\mathfrak{a}_n) \in \mathbb{D}_0$. When $0 \leqslant r \leqslant k$, we define the mean value $K_{a,b}^r = K_{a,b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a})$ by putting

$$K_{a,b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a}) = \rho_0^{-4} \sum_{\xi \bmod p^a} \sum_{\substack{\eta \bmod p^b \\ \xi \not\equiv \eta \pmod{p^\nu}}} \rho_a(\xi)^2 \rho_b(\eta)^2 K_{a,b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a};\xi,\eta), \tag{3.19}$$

in which

$$K_{a,b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a};\xi,\eta) = \oint_{p^B} |\mathfrak{f}_a(\boldsymbol{\alpha};\xi)^{2R} \mathfrak{f}_b(\boldsymbol{\alpha};\eta)^{2s-2R}| \, \mathrm{d}\boldsymbol{\alpha} \tag{3.20}$$

and $R = r(r+1)/2$. Notice here that in the definition (3.20) of $K_{a,b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a};\xi,\eta)$, the parameter ν is in a sense otiose. However, its inclusion as a superscript serves to remind us of the implicit assumption that $p^\nu \nmid (\xi - \eta)$. By orthogonality, the mean value $K_{a,b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a};\xi,\eta)$ counts the integral solutions of the simultaneous congruences

$$\sum_{i=1}^{R} (\varphi_j(x_i) - \varphi_j(y_i)) \equiv \sum_{l=1}^{s-R} (\varphi_j(v_l) - \varphi_j(w_l)) \pmod{p^B} \quad (1 \leqslant j \leqslant k), \tag{3.21}$$

satisfying

$$\mathbf{x} \equiv \mathbf{y} \equiv \xi \pmod{p^a} \quad \text{and} \quad \mathbf{v} \equiv \mathbf{w} \equiv \eta \pmod{p^b}, \tag{3.22}$$

with each solution being counted with weight

$$\rho_a(\xi)^{-2R} \rho_b(\eta)^{2R-2s} \Big(\prod_{i=1}^{R} \mathfrak{a}_{x_i} \overline{\mathfrak{a}}_{y_i} \Big) \Big(\prod_{l=1}^{s-R} \mathfrak{a}_{v_l} \overline{\mathfrak{a}}_{w_l} \Big). \tag{3.23}$$

As in our previous work on efficient congruencing, our arguments are considerably simplified by making transparent the relationship between various mean values, on the one hand, and their anticipated magnitudes, on the other.

We therefore consider normalised versions of these mean values $K_{a,b}^r$ as follows. When $1 \leqslant r \leqslant k-1$ and Δ is a positive number, we define

$$[\![K_{a,b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a})]\!]_\Delta = \left(\frac{K_{a,b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a})}{p^{\Delta H} U_{s,k}^{B,H}(\mathfrak{a})} \right)^{\frac{k-1}{r(k-r)}}. \tag{3.24}$$

For much of our discussion, the choices of $\mathfrak{a}$, $\boldsymbol{\varphi}$, ν and c will be considered fixed, and in such circumstances we suppress mention of them from our notation. We note that the presence of the exponent

$$\frac{k-1}{r(k-r)}$$

is designed to equalise the weights with which the mixed mean values $K_{a,b}^r$ occur within our arguments. The utility of this device will become apparent in due course.

4. Translation-dilation invariant families

Consider a p^c-spaced system of polynomials $\boldsymbol{\varphi}$. In general, of course, the system of congruences

$$\sum_{i=1}^{s} \varphi_j(x_i) \equiv \sum_{i=1}^{s} \varphi_j(y_i) \pmod{p^B} \quad (1 \leqslant j \leqslant k) \tag{4.1}$$

will not be translation-dilation invariant. However, in such special cases as the system given by $\varphi_j(t) = t^j$ $(1 \leqslant j \leqslant k)$, it follows via an application of the binomial theorem that whenever $a \in \mathbb{Z}$ and $q \in \mathbb{N}$, and $(\mathbf{x}, \mathbf{y}) = (q\mathbf{u}+a, q\mathbf{v}+a)$ is a solution of (4.1), then so too is $(\mathbf{x}, \mathbf{y}) = (\mathbf{u}, \mathbf{v})$. This translation-dilation invariance property may be preserved in *families* of systems $\boldsymbol{\varphi}$ which are not individually translation-dilation invariant. Thus, given $a \in \mathbb{Z}$ and $b \in \mathbb{N}$, and a solution $(\mathbf{x}, \mathbf{y}) = (p^b\mathbf{u} + a, p^b\mathbf{v} + a)$ of the system (4.1), one finds that $(\mathbf{u}, \mathbf{v})$ is a solution of the system

$$\sum_{i=1}^{s} \psi_j(u_i) \equiv \sum_{i=1}^{s} \psi_j(v_i) \pmod{p^{B'}} \quad (1 \leqslant j \leqslant k),$$

for some other p^c-spaced system of polynomials $\boldsymbol{\psi}$, with B' an integer depending on B and b. It is our goal in this section to establish estimates making this property explicit.

Our first lemma shows that restriction to arithmetic progressions modulo p^h leads to bounds on mean values of the type $U_{s,k}^B(\mathfrak{a})$ that scale appropriately with respect to the height of the arithmetic progression.

Lemma 4.1. *Suppose that $k \in \mathbb{N}$ and that p is a prime number with $p > k$. Let τ, ε and δ be positive numbers with $\varepsilon < \tau < \delta < 1$, and let B be sufficiently large in terms of s, k and ε. Write $H = \lceil B/k \rceil$. Then for every $\boldsymbol{\varphi} \in \Phi_\tau(B)$, every sequence $(\mathfrak{a}_n) \in \mathbb{D}_0$, and every non-negative integer h with $h \leqslant (1-\delta)H$, one has*

$$U_{s,k}^{B,h}(\mathfrak{a}) \ll (p^{H-h})^{\lambda(s,k)+\varepsilon} U_{s,k}^{B,H}(\mathfrak{a}).$$

Proof. Consider a p^c-spaced k-tuple of polynomials $\boldsymbol{\varphi}$ with $c \geqslant \tau B$, and a complex sequence $(\mathfrak{a}_n) \in \mathbb{D}_0$, with associated parameters satisfying the hypotheses of the statement of the lemma. Also, let h be an integer with $0 \leqslant h \leqslant (1-\delta)H$. From the definition (3.4) of the exponential sum $\mathfrak{f}_h(\boldsymbol{\alpha};\xi)$, one has

$$\mathfrak{f}_h(\boldsymbol{\alpha};\xi) = \rho_h(\xi)^{-1} \sum_{y \in \mathbb{Z}} \mathfrak{b}_y(\xi) e(\psi(p^h y + \xi; \boldsymbol{\alpha})), \tag{4.2}$$

in which $\psi(n;\boldsymbol{\alpha})$ is given by (3.3), and the coefficients $\mathfrak{b}_y = \mathfrak{b}_y(\xi)$ are defined by putting

$$\mathfrak{b}_y(\xi) = \mathfrak{a}_{p^h y + \xi} \quad (y \in \mathbb{Z}). \tag{4.3}$$

By orthogonality, the integral on the right hand side of (3.8) counts the integral solutions $\mathbf{y}, \mathbf{z}$ of the simultaneous congruences

$$\sum_{i=1}^{s} \varphi_j(p^h y_i + \xi) \equiv \sum_{i=1}^{s} \varphi_j(p^h z_i + \xi) \pmod{p^B} \quad (1 \leqslant j \leqslant k), \tag{4.4}$$

with each solution being counted with weight

$$\rho_h(\xi)^{-2s} \prod_{i=1}^{s} \mathfrak{b}_{y_i} \overline{\mathfrak{b}}_{z_i}. \tag{4.5}$$

The polynomial system $\boldsymbol{\varphi}$ is p^c-spaced, and hence it follows from (3.1) that for suitable polynomials $\psi_j \in \mathbb{Z}[t]$, one may write

$$\varphi_j(t) = \sum_{i=1}^{k} \omega_{ij} t^i + p^c t^{k+1} \psi_j(t) \quad (1 \leqslant j \leqslant k),$$

for some integral coefficient matrix $A_1 = (\omega_{ij})_{1 \leqslant i,j \leqslant k}$ congruent modulo p^c to the $k \times k$ identity matrix I_k. Since $A_1 \equiv I_k \pmod{p^c}$, it follows that A_1 possesses a multiplicative inverse A_1^{-1} modulo p^B having integral coefficients. By replacing $\boldsymbol{\varphi}$ by $A_1^{-1}\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$ by $A_1^{-1}\boldsymbol{\psi}$, which amounts to taking suitable integral linear combinations of the congruences comprising (4.4), we discern that there is no loss of generality in supposing that for $1 \leqslant j \leqslant k$, one has

$$\varphi_j(t) = t^j + p^c t^{k+1} \psi_j(t) \quad (1 \leqslant j \leqslant k). \tag{4.6}$$

With the latter assumption in hand, we apply the binomial theorem to (4.6). Thus, when $1 \leqslant j \leqslant k$, one finds that for suitable polynomials $\Psi_j \in \mathbb{Z}[t]$, one has $\varphi_j(p^h y + \xi) - \varphi_j(\xi) = \Phi_j(p^h y)$, in which

$$\Phi_j(t) = \sum_{i=1}^{k} \Omega_{ij} \xi^{j-i} t^i + p^c t^{k+1} \Psi_j(t) \quad (1 \leqslant j \leqslant k),$$

and the integral coefficients Ω_{ij} satisfy

$$\Omega_{ij} \equiv \binom{j}{i} \pmod{p^c} \quad (1 \leqslant i, j \leqslant k).$$

Here, we adopt the convention that the binomial coefficient $\binom{j}{i}$ is zero for $i > j$. Our hypothesis that $p > k$ ensures that the matrix $A_2 = (\Omega_{ij})_{1 \leqslant i,j \leqslant k}$ possesses a multiplicative inverse A_2^{-1} modulo p^B having integral coefficients,

since it is triangular modulo p^c with diagonal entries all equal to 1. We now replace $\mathbf{\Phi}$ by $A_2^{-1}\mathbf{\Phi}$ and $\mathbf{\Psi}$ by $A_2^{-1}\mathbf{\Psi}$. Again, this amounts to taking suitable integral linear combinations of the congruences comprising (4.4), and we see that there is no loss of generality in supposing that the coefficient matrix A_2 is equal to I_k. Hence, there exist polynomials $\Upsilon_j \in \mathbb{Z}[t]$ having the property that whenever the system (4.4) is satisfied, then

$$\sum_{i=1}^{s}(p^h)^j\,(\Phi_j(y_i)-\Phi_j(z_i)) \equiv 0 \pmod{p^B} \quad (1 \leqslant j \leqslant k), \tag{4.7}$$

in which Φ_j has been redefined by

$$\Phi_j(t) = t^j + p^{c+h}t^{k+1}\Upsilon_j(t). \tag{4.8}$$

The integral solutions $\mathbf{y}, \mathbf{z}$ of the original system of congruences (4.4), counted with the weight (4.5) associated with the definition (3.8) of $U_{s,k}^{B,h}(\mathfrak{a})$, are therefore constrained by the additional system of congruences

$$\sum_{i=1}^{s}\Phi_j(y_i) \equiv \sum_{i=1}^{s}\Phi_j(z_i) \pmod{p^{B-kh}} \quad (1 \leqslant j \leqslant k). \tag{4.9}$$

In order to incorporate the extra condition (4.9) into the mean value $U_{s,k}^{B,h}(\mathfrak{a})$, we introduce the exponential sum

$$\mathfrak{g}_h(\boldsymbol{\alpha},\boldsymbol{\beta};\xi) = \rho_h(\xi)^{-1}\sum_{y\in\mathbb{Z}}\mathfrak{c}_y(\xi;\boldsymbol{\alpha})e(\psi^*(y;\boldsymbol{\beta})), \tag{4.10}$$

where

$$\mathfrak{c}_y(\xi;\boldsymbol{\alpha}) = \mathfrak{b}_y(\xi)e(\psi(p^hy+\xi;\boldsymbol{\alpha})) \tag{4.11}$$

and

$$\psi^*(y;\boldsymbol{\beta}) = \beta_1\Phi_1(y) + \ldots + \beta_k\Phi_k(y).$$

Equipped with this notation, it follows via orthogonality that

$$\oint_{p^B}|\mathfrak{f}_h(\boldsymbol{\alpha};\xi)|^{2s}\,\mathrm{d}\boldsymbol{\alpha} = \oint_{p^B}\oint_{p^{B-kh}}|\mathfrak{g}_h(\boldsymbol{\alpha},\boldsymbol{\beta};\xi)|^{2s}\,\mathrm{d}\boldsymbol{\beta}\,\mathrm{d}\boldsymbol{\alpha}. \tag{4.12}$$

Note that the polynomial system $\mathbf{\Phi}$ defined via (4.8) is p^{c+h}-spaced. Since we may assume that $h \leqslant (1-\delta)H$, moreover, one has

$$B - kh \geqslant B - k\lceil B/k\rceil + k\delta\lceil B/k\rceil \geqslant \delta B - k,$$

so that $B-kh$ may be assumed to be sufficiently large in terms of s, k and ε. It therefore follows from the definitions (3.12) and (3.13) that

$$U_{s,k}^{B-kh}(\mathfrak{c}) \ll (p^{H-h})^{\lambda(s,k)+\varepsilon}U_{s,k}^{B-kh,H-h}(\mathfrak{c}).$$

In view of the definition (4.10), we therefore deduce that

$$\oint_{p^{B-kh}}|\mathfrak{g}_h(\boldsymbol{\alpha},\boldsymbol{\beta};\xi)|^{2s}\,\mathrm{d}\boldsymbol{\beta} \ll (p^{H-h})^{\lambda(s,k)+\varepsilon}I_1, \tag{4.13}$$

where

$$I_1 = \rho_h(\xi)^{-2} \sum_{\eta \bmod p^{H-h}} \rho_{H-h}(p^h\eta+\xi)^2 \oint_{p^{B-kh}} |\mathfrak{g}^*_{H-h}(\boldsymbol{\alpha},\boldsymbol{\beta};\xi,\eta)|^{2s}\,\mathrm{d}\boldsymbol{\beta},$$

and

$$\mathfrak{g}^*_{H-h}(\boldsymbol{\alpha},\boldsymbol{\beta};\xi,\eta) = \rho_{H-h}(p^h\eta+\xi)^{-1} \sum_{y\equiv\eta\ (\mathrm{mod}\ p^{H-h})} \mathfrak{c}_y(\xi;\boldsymbol{\alpha})e(\psi^*(y;\boldsymbol{\beta})).$$

Observe that there is a correspondence between residues ζ modulo p^H and pairs (ξ,η) of residues modulo p^h, and modulo p^{H-h}, respectively. Indeed, if $1 \leqslant \zeta \leqslant p^H$, then we may identify ξ and η via the relations $\xi \equiv \zeta\ (\mathrm{mod}\ p^h)$ and $\eta \equiv (\zeta-\xi)p^{-h}\ (\mathrm{mod}\ p^{H-h})$. Likewise, given (ξ,η), we put $\zeta = p^h\eta+\xi$. With this correspondence in mind, we find from (4.3) and (4.11) that $\mathfrak{g}^*_{H-h}(\boldsymbol{\alpha},\boldsymbol{\beta};\xi,\eta)$ is equal to

$$\rho_H(\zeta)^{-1} \sum_{\substack{y\in\mathbb{Z}\\ p^hy+\xi\equiv\zeta\ (\mathrm{mod}\ p^H)}} \mathfrak{a}_{p^hy+\xi} e(\psi(p^hy+\xi;\boldsymbol{\alpha})+\psi^*(y;\boldsymbol{\beta})).$$

In particular, we deduce from (4.12) and (4.13) that

$$\begin{aligned}\sum_{\xi \bmod p^h} \rho_h(\xi)^2 \oint_{p^B} |\mathfrak{f}_h(\boldsymbol{\alpha};\xi)|^{2s}\,\mathrm{d}\boldsymbol{\alpha}&\\ \ll (p^{H-h})^{\lambda(s,k)+\varepsilon} \sum_{\substack{\zeta \bmod p^H\\ \zeta=p^h\eta+\xi}} \rho_H(\zeta)^2 \oint_{p^B}\oint_{p^{B-kh}} |\mathfrak{g}^*_{H-h}(\boldsymbol{\alpha},\boldsymbol{\beta};\xi,\eta)|^{2s}\,\mathrm{d}\boldsymbol{\beta}\,\mathrm{d}\boldsymbol{\alpha}.&\end{aligned} \tag{4.14}$$

By orthogonality, the mean value on the right hand side of (4.14) counts integral solutions of the system of congruences (4.4), with each solution being counted with weight (4.5), but subject to the additional condition that

$$p^h\mathbf{y}+\xi \equiv p^h\mathbf{z}+\xi \equiv \zeta\ (\mathrm{mod}\ p^H),$$

and further subject to the congruence conditions (4.9). The latter congruence conditions are generated by the integral over $\boldsymbol{\beta}$ in (4.14). However, the conditions (4.9) are implied by (4.4), as we have shown in the discussion above. We note also that when $\zeta = p^h\eta+\xi$, then

$$\mathfrak{g}^*_{H-h}(\boldsymbol{\alpha},\mathbf{0};\xi,\eta) = \rho_H(\zeta)^{-1} \sum_{n\equiv\zeta\ (\mathrm{mod}\ p^H)} \mathfrak{a}_n e(\psi(n;\boldsymbol{\alpha})) = \mathfrak{f}_H(\boldsymbol{\alpha};\zeta).$$

Then on recalling the definition (3.8), and noting that the conditions (4.9) may be omitted, we deduce that

$$\begin{aligned}U^{B,h}_{s,k}(\mathfrak{a}) &\ll (p^{H-h})^{\lambda(s,k)+\varepsilon}\rho_0^{-2} \sum_{\zeta \bmod p^H} \rho_H(\zeta)^2 \oint_{p^B} |\mathfrak{f}_H(\boldsymbol{\alpha};\zeta)|^{2s}\,\mathrm{d}\boldsymbol{\alpha}\\ &= (p^{H-h})^{\lambda(s,k)+\varepsilon} U^{B,H}_{s,k}(\mathfrak{a}).\end{aligned}$$

This completes the proof of the lemma. □

It is tempting to replace the proof of Lemma 4.1 with an informal argument appealing to Taylor expansions, rescaling and an appeal to translation invariance. However, as should be apparent from our proof above, there are subtle technical issues that arise in a detailed argument that threaten to sabotage the desired conclusion. We therefore offer no apology (beyond this remark) for expending space on the detailed account above.

The estimate supplied by Lemma 4.1 is easily transformed into a crude estimate for the mean value $K_{a,b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a})$ which nonetheless has considerable utility.

Lemma 4.2. *Suppose that $k \in \mathbb{N}$ and that p is a prime number with $p > k$. Let τ, ε and δ be positive numbers with $\varepsilon < \tau < \delta < 1$, and let B be sufficiently large in terms of s, k and ε. Write $H = \lceil B/k \rceil$ and suppose that r and ν are non-negative integers with $1 \leqslant r \leqslant k-1$. In addition, suppose that $0 < \Lambda \leqslant \lambda(s,k)$. Then for every $\boldsymbol{\varphi} \in \Phi_\tau(B)$, every sequence $(\mathfrak{a}_n) \in \mathbb{D}_0$, and all non-negative integers a and b satisfying $\max\{a,b\} \leqslant (1-\delta)H$, one has*

$$[\![K_{a,b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a})]\!]_\Lambda \ll (p^H)^{\lambda(s,k)-\Lambda+\varepsilon}.$$

Proof. An application of Hölder's inequality delivers the estimate

$$\rho_0^{-4} \sum_{\xi \bmod p^a} \sum_{\eta \bmod p^b} \rho_a(\xi)^2 \rho_b(\eta)^2 \oint_{p^B} |\mathfrak{f}_a(\boldsymbol{\alpha};\xi)^{2R} \mathfrak{f}_b(\boldsymbol{\alpha};\eta)^{2s-2R}| \, \mathrm{d}\boldsymbol{\alpha} \leqslant I_1^{R/s} I_2^{1-R/s}, \tag{4.15}$$

where

$$I_1 = \rho_0^{-4} \sum_{\xi \bmod p^a} \sum_{\eta \bmod p^b} \rho_a(\xi)^2 \rho_b(\eta)^2 \oint_{p^B} |\mathfrak{f}_a(\boldsymbol{\alpha};\xi)|^{2s} \, \mathrm{d}\boldsymbol{\alpha}$$

and

$$I_2 = \rho_0^{-4} \sum_{\xi \bmod p^a} \sum_{\eta \bmod p^b} \rho_a(\xi)^2 \rho_b(\eta)^2 \oint_{p^B} |\mathfrak{f}_b(\boldsymbol{\alpha};\eta)|^{2s} \, \mathrm{d}\boldsymbol{\alpha}.$$

By reference to (3.2) and (3.8), one finds that

$$I_1 = \rho_0^{-2} \bigg(\sum_{\eta \bmod p^b} \rho_b(\eta)^2 \bigg) U_{s,k}^{B,a}(\mathfrak{a}) = U_{s,k}^{B,a}(\mathfrak{a}),$$

and likewise one sees that $I_2 = U_{s,k}^{B,b}(\mathfrak{a})$. Consequently, on applying Lemma 4.1, we deduce from (3.19) and (4.15) that

$$K_{a,b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a}) \ll \big((p^{H-a})^{\lambda(s,k)+\varepsilon}\big)^{R/s} \big((p^{H-b})^{\lambda(s,k)+\varepsilon}\big)^{1-R/s} U_{s,k}^{B,H}(\mathfrak{a}). \tag{4.16}$$

Next we recall (3.24). This conveys us from the estimate (4.16) to the corresponding normalised bound

$$[\![K_{a,b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a})]\!]_\Lambda^{\frac{r(k-r)}{k-1}} \ll \big((p^{H-a})^{R/s} (p^{H-b})^{1-R/s}\big)^{\lambda(s,k)+\varepsilon} p^{-\Lambda H}.$$

But when $1 \leqslant r \leqslant k-1$, one finds that $r(k-r) \geqslant k-1$, whence

$$\frac{r(k-r)}{k-1} \geqslant 1.$$

Thus we conclude that

$$[\![K_{a,b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a})]\!]_\Lambda \ll \big(p^{H(\lambda(s,k)-\Lambda+\varepsilon)}\big)^{\frac{k-1}{r(k-r)}} \ll (p^H)^{\lambda(s,k)-\Lambda+\varepsilon}.$$

This completes the proof of the lemma. □

5. The base of the induction: the trivial case $k = 1$

Most of the basic infrastructure and the skeletal properties of the key mean values are now in place. Our overarching strategy is to establish Theorem 3.1 by induction. Thus, assuming the validity of Theorem 3.1 for exponents smaller than k, we seek to establish its conclusion for the exponent k. We begin in this section by establishing the base case $k = 1$. Although essentially trivial, the underlying ideas are instructive. We remark that the passing similarity with the argument underlying the conventional proof of Hensel's lemma is not accidental.

Lemma 5.1. *For any prime number p, one has $\lambda(1, 1) = 0$.*

Proof. Let τ be a small positive number, and let B be a positive integer sufficiently large in terms of τ. Consider any sequence $(\mathfrak{a}_n) \in \mathbb{D}$ and polynomial $\varphi \in \Phi_\tau(B)$. We may suppose that φ is a p^c-spaced polynomial for some $c \geqslant \tau B$, so that $\varphi(t) = t + p^c\psi(t)$ for a polynomial $\psi \in \mathbb{Z}[t]$. The mean value $U_{1,1}^B(\mathfrak{a})$ counts the integral solutions of the congruence

$$\varphi(x) \equiv \varphi(y) \pmod{p^B}, \tag{5.1}$$

with each solution x, y being counted with weight $\rho_0^{-2}\mathfrak{a}_x\overline{\mathfrak{a}}_y$. The congruence (5.1) is in fact the relation

$$x + p^c\psi(x) \equiv y + p^c\psi(y) \pmod{p^B}, \tag{5.2}$$

from which we infer that $x \equiv y \pmod{(p^c, p^B)}$. Here, we take the liberty of writing $p^{\min\{c,B\}}$ as the highest common factor (p^c, p^B), this being cosmetically slightly less awkward.

Since $x - y$ divides $\psi(x) - \psi(y)$, it follows that the highest common factor (p^c, p^B) divides $\psi(x) - \psi(y)$. Substituting this relation back into (5.2), we deduce that $x \equiv y \pmod{(p^{2c}, p^B)}$. By repeating this argument no more than $\lceil 1/\tau \rceil$ times, we conclude that $x \equiv y \pmod{p^B}$. We may therefore classify the solutions of the congruence (5.2) according to the common congruence class ξ modulo p^B of x and y. On recalling the definitions (3.4), (3.6) and (3.8), we infer via orthogonality that

$$U_{1,1}^B(\mathfrak{a}) = \rho_0^{-2} \sum_{1 \leqslant \xi \leqslant p^B} \rho_B(\xi)^2 \oint_{p^B} |\mathfrak{f}_B(\boldsymbol{\alpha}; \xi)|^2 \, \mathrm{d}\boldsymbol{\alpha} = U_{1,1}^{B,B}(\mathfrak{a}).$$

Thus we conclude that

$$\log\Big(U_{1,1}^B(\mathfrak{a})/U_{1,1}^{B,B}(\mathfrak{a})\Big) = 0,$$

and it is immediate from (3.12) that $\lambda^*(1, 1; \tau) = 0$. We therefore conclude from (3.13) that $\lambda(1, 1) = 0$, completing the proof of the lemma. □

6. The initial conditioning process

We now move on to our main inductive task of establishing Theorem 3.1 for the exponent $k \geqslant 2$, assuming its validity for exponents smaller than k. The initial step in the estimation of $U_{s,k}^B(\mathfrak{a})$ is to bound it in terms of a mean value of the shape $K_{a,b,c}^{1,\boldsymbol{\varphi},\nu}(\mathfrak{a})$. In this section we describe the initial set-up to be applied in later sections, as well as the conditioning of the variables underlying the mean value $K_{a,b,c}^{1,\boldsymbol{\varphi},\nu}(\mathfrak{a})$.

Recall the definitions (3.12) and (3.13). We fix a prime number p with $p > k$ and we also fix $s = k(k+1)/2$, and then we seek to show that $\lambda(s,k) = 0$. This we achieve by deriving a contradiction to the assumption that $\lambda(s,k) = \Lambda > 0$. Note here that in view of (3.14), there is no loss of generality in assuming that

$$0 < \Lambda \leqslant s. \tag{6.1}$$

We next introduce a hierarchy of sufficiently small positive numbers ε, τ, δ and μ with

$$\varepsilon < \tau < \delta < \mu < 1. \tag{6.2}$$

We suppose that each element in the hierarchy is sufficiently small in terms of k, Λ, and the larger elements of the hierarchy. It follows from (3.13) that there is no loss of generality in supposing that $\lambda^*(s,k;\tau) \geqslant \Lambda - \varepsilon/2$. Thus, in view of (3.12), there exists a sequence $(B_m)_{m=1}^\infty$, with $B_m \to \infty$ as $m \to \infty$, having the property that whenever m is sufficiently large, then there exists $\boldsymbol{\varphi}_m \in \Phi_\tau(B_m)$ and $(\mathfrak{a}_n^{(m)}) \in \mathbb{D}$ for which

$$U_{s,k}^{B_m}(\mathfrak{a}^{(m)}) \geqslant (p^{H_m})^{\Lambda-\varepsilon} U_{s,k}^{B_m,H_m}(\mathfrak{a}^{(m)}), \tag{6.3}$$

where we have written $H_m = \lceil B_m/k \rceil$. We now fix such an integer m with the property that $B = B_m$ is sufficiently large in terms of k, Λ, μ, δ, τ and ε. In our discussion to come, we may henceforth omit mention of the subscript or superscript m in $\boldsymbol{\varphi}_m$, $(\mathfrak{a}_n^{(m)})$, H_m. Observe that since B is sufficiently large in terms of the basic parameters, then we may suppose also that εB is also sufficiently large in terms of k, Λ, μ, τ, δ and ε.

Also, by virtue of Lemma 4.1, we may assume that for every non-negative integer h with $h \leqslant (1-\delta)H$, uniformly in $\boldsymbol{\varphi}' \in \Phi_\tau(B)$ and $(\mathfrak{a}'_n) \in \mathbb{D}_0$, one has

$$U_{s,k}^{B,h}(\mathfrak{a}') \leqslant (p^{H-h})^{\Lambda+\varepsilon} U_{s,k}^{B,H}(\mathfrak{a}'). \tag{6.4}$$

We may now announce our first bound for $U_{s,k}^B(\mathfrak{a})$ in terms of mean values of the shape $K_{a,b,c}^{1,\boldsymbol{\varphi},\nu}(\mathfrak{a})$. Here, we fix a choice for the parameter ν for future use by setting

$$\nu = \lceil 4\varepsilon H \Lambda^{-1} \rceil. \tag{6.5}$$

Since we may suppose that $\boldsymbol{\varphi} \in \Phi_\tau(B)$, it follows that there exists an integer c with $c \geqslant \tau B$ with the property that $\boldsymbol{\varphi}$ is p^c-spaced.

Lemma 6.1. *One has* $U_{s,k}^B(\mathfrak{a}) \ll p^{s\nu} K_{\nu,\nu,c}^{1,\boldsymbol{\varphi},\nu}(\mathfrak{a})$.

Proof. In view of the definition of $f_{\mathfrak{a}}(\boldsymbol{\alpha})$ and the definition (3.4), one has

$$\rho_0^2 f_{\mathfrak{a}}(\boldsymbol{\alpha})^2 = \sum_{\xi \bmod p^\nu} \rho_\nu(\xi) \mathfrak{f}_\nu(\boldsymbol{\alpha};\xi) \left(\rho_0 f_{\mathfrak{a}}(\boldsymbol{\alpha})\right).$$

Moreover, for each given residue ξ modulo p^ν, one has

$$\rho_0 f_{\mathfrak{a}}(\boldsymbol{\alpha}) = \rho_\nu(\xi) f_\nu(\boldsymbol{\alpha};\xi) + \sum_{\substack{\eta \bmod p^\nu \\ \eta \not\equiv \xi \pmod{p^\nu}}} \rho_\nu(\eta) f_\nu(\boldsymbol{\alpha};\eta).$$

Thus

$$|f_{\mathfrak{a}}(\boldsymbol{\alpha})|^2 \leqslant T_1(\boldsymbol{\alpha}) + T_2(\boldsymbol{\alpha}), \tag{6.6}$$

where

$$T_1(\boldsymbol{\alpha}) = \rho_0^{-2} \sum_{\xi \bmod p^\nu} \rho_\nu(\xi)^2 |f_\nu(\boldsymbol{\alpha};\xi)|^2$$

and

$$T_2(\boldsymbol{\alpha}) = \rho_0^{-2} \sum_{\xi \bmod p^\nu} \sum_{\substack{\eta \bmod p^\nu \\ \eta \not\equiv \xi \pmod{p^\nu}}} \rho_\nu(\xi)\rho_\nu(\eta) |f_\nu(\boldsymbol{\alpha};\xi) f_\nu(\boldsymbol{\alpha};\eta)|.$$

In much the same manner as in the derivation of the relation (3.9), an application of Hölder's inequality reveals that

$$\begin{aligned} T_1(\boldsymbol{\alpha})^s &\leqslant \rho_0^{-2s} \bigg(\sum_{\xi \bmod p^\nu} \rho_\nu(\xi)^2 \bigg)^{s-1} \sum_{\xi \bmod p^\nu} \rho_\nu(\xi)^2 |\mathfrak{f}_\nu(\boldsymbol{\alpha};\xi)|^{2s} \\ &= \rho_0^{-2} \sum_{\xi \bmod p^\nu} \rho_\nu(\xi)^2 |\mathfrak{f}_\nu(\boldsymbol{\alpha};\xi)|^{2s}. \end{aligned} \tag{6.7}$$

Meanwhile, again applying Hölder's inequality, one finds that

$$T_2(\boldsymbol{\alpha})^{2s} \leqslant \rho_0^{-4s} T_3^s T_4^{s-2} T_5 T_6,$$

where

$$\begin{aligned} T_3 &= \sum_{\xi \bmod p^\nu} \sum_{\eta \bmod p^\nu} 1, \\ T_4 &= \sum_{\xi \bmod p^\nu} \sum_{\eta \bmod p^\nu} \rho_\nu(\xi)^2 \rho_\nu(\eta)^2, \\ T_5 &= \sum_{\xi \bmod p^\nu} \sum_{\substack{\eta \bmod p^\nu \\ \eta \not\equiv \xi \pmod{p^\nu}}} \rho_\nu(\xi)^2 \rho_\nu(\eta)^2 |\mathfrak{f}_\nu(\boldsymbol{\alpha};\xi)^2 \mathfrak{f}_\nu(\boldsymbol{\alpha};\eta)^{2s-2}|, \\ T_6 &= \sum_{\xi \bmod p^\nu} \sum_{\substack{\eta \bmod p^\nu \\ \eta \not\equiv \xi \pmod{p^\nu}}} \rho_\nu(\xi)^2 \rho_\nu(\eta)^2 |\mathfrak{f}_\nu(\boldsymbol{\alpha};\eta)^2 \mathfrak{f}_\nu(\boldsymbol{\alpha};\xi)^{2s-2}|. \end{aligned}$$

We have $T_3 = p^{2\nu}$ and $T_4 = \rho_0^4$. Also, by symmetry, one has $T_5 = T_6$. We therefore deduce that

$$T_2(\boldsymbol{\alpha})^s \leqslant \rho_0^{-4} p^{\nu s} \sum_{\xi \bmod p^\nu} \sum_{\substack{\eta \bmod p^\nu \\ \eta \not\equiv \xi \pmod{p^\nu}}} \rho_\nu(\xi)^2 \rho_\nu(\eta)^2 |\mathfrak{f}_\nu(\boldsymbol{\alpha};\xi)^2 \mathfrak{f}_\nu(\boldsymbol{\alpha};\eta)^{2s-2}|. \tag{6.8}$$

On recalling the definitions (3.8) and (3.19), we deduce from (6.7) and (6.8) that

$$\oint_{p^B} T_1(\boldsymbol{\alpha})^s \,\mathrm{d}\boldsymbol{\alpha} \leqslant U_{s,k}^{B,\nu}(\mathfrak{a}) \quad \text{and} \quad \oint_{p^B} T_2(\boldsymbol{\alpha})^s \,\mathrm{d}\boldsymbol{\alpha} \leqslant p^{\nu s} K_{\nu,\nu,c}^{1,\boldsymbol{\varphi},\nu}(\mathfrak{a}).$$

Since it follows from (6.6) that

$$|f_{\mathfrak{a}}(\boldsymbol{\alpha})|^{2s} \ll |T_1(\boldsymbol{\alpha})|^s + |T_2(\boldsymbol{\alpha})|^s,$$

we deduce from the definition (3.6) that

$$U_{s,k}^{B}(\mathfrak{a}) \ll U_{s,k}^{B,\nu}(\mathfrak{a}) + p^{\nu s} K_{\nu,\nu,c}^{1,\boldsymbol{\varphi},\nu}(\mathfrak{a}). \tag{6.9}$$

Next we make use of the estimate (6.4) to obtain the bound

$$U_{s,k}^{B,\nu}(\mathfrak{a}) \leqslant (p^{H-\nu})^{\Lambda+\varepsilon} U_{s,k}^{B,H}(\mathfrak{a}).$$

Since the definition (6.5) ensures that

$$(\Lambda+\varepsilon)(H-\nu) - (\Lambda-\varepsilon)H = 2\varepsilon H - (\Lambda+\varepsilon)\nu < -2\varepsilon H,$$

we conclude via (6.3) that

$$U_{s,k}^{B,\nu}(\mathfrak{a}) \leqslant p^{-2\varepsilon H}(p^H)^{\Lambda-\varepsilon} U_{s,k}^{B,H}(\mathfrak{a}) \leqslant p^{-2\varepsilon H} U_{s,k}^{B}(\mathfrak{a}).$$

Thus we infer from (6.9) that

$$U_{s,k}^{B}(\mathfrak{a}) \ll p^{-2\varepsilon H} U_{s,k}^{B}(\mathfrak{a}) + p^{\nu s} K_{\nu,\nu,c}^{1,\boldsymbol{\varphi},\nu}(\mathfrak{a}),$$

whence

$$U_{s,k}^{B}(\mathfrak{a}) \ll p^{\nu s} K_{\nu,\nu,c}^{1,\boldsymbol{\varphi},\nu}(\mathfrak{a}).$$

This completes the proof of the lemma. □

The congruence condition implicit in the mean value $K_{\nu,\nu,c}^{1,\boldsymbol{\varphi},\nu}(\mathfrak{a})$ is not yet strong enough to pursue our main iteration, and so we pause to strengthen it before proceeding further. It is at this point that the parameter μ enters the scene. We put

$$\theta = \lceil \mu H \rceil. \tag{6.10}$$

We briefly pause to record a useful, though essentially trivial, upper bound permitting us to relate exponential sums restricted to congruences associated with differing powers of p.

Lemma 6.2. *Suppose that a and b are non-negative integers with $a \leqslant b$. Then for any positive number w and any integer ξ, one has*

$$\rho_a(\xi)^2 |\mathfrak{f}_a(\boldsymbol{\alpha};\xi)|^{2w} \leqslant (p^{b-a})^w \sum_{\substack{\zeta \bmod p^b \\ \zeta \equiv \xi \ (\mathrm{mod}\ p^a)}} \rho_b(\zeta)^2 |\mathfrak{f}_b(\boldsymbol{\alpha};\zeta)|^{2w}.$$

Proof. By applying Hölder's inequality in combination with (3.4), one obtains

$$\rho_a(\xi)^{2w}|f_a(\boldsymbol{\alpha};\xi)|^{2w} = \Big| \sum_{\substack{\zeta \bmod p^b \\ \zeta \equiv \xi \ (\mathrm{mod}\ p^a)}} \rho_b(\zeta)\mathfrak{f}_b(\boldsymbol{\alpha};\zeta)\Big|^{2w}$$
$$\leqslant U_1^w U_2^{w-1} \sum_{\substack{\zeta \bmod p^b \\ \zeta \equiv \xi \ (\mathrm{mod}\ p^a)}} \rho_b(\zeta)^2 |\mathfrak{f}_b(\boldsymbol{\alpha};\zeta)|^{2w}, \tag{6.11}$$

where

$$U_1 = \sum_{\substack{\zeta \bmod p^b \\ \zeta \equiv \xi \ (\mathrm{mod}\ p^a)}} 1 = p^{b-a}$$

and

$$U_2 = \sum_{\substack{\zeta \bmod p^b \\ \zeta \equiv \xi \ (\mathrm{mod}\ p^a)}} \rho_b(\zeta)^2 = \sum_{n \equiv \xi \ (\mathrm{mod}\ p^a)} |\mathfrak{a}_n|^2 = \rho_a(\xi)^2.$$

The desired conclusion is immediate on substituting the latter relations into (6.11) □

Lemma 6.3. *One has* $U_{s,k}^B(\mathfrak{a}) \ll p^{s\theta} K_{\theta,\theta,c}^{1,\boldsymbol{\varphi},\nu}(\mathfrak{a})$.

Proof. On recalling (3.20), we deduce via two applications of Lemma 6.2 that the mean value $K_{\nu,\nu,c}^{1,\boldsymbol{\varphi},\nu}(\mathfrak{a};\xi,\eta)$ is bounded above by

$$\rho_\nu(\xi)^{-2}\rho_\nu(\eta)^{-2}(p^{\theta-\nu})^s \sum_{\xi',\eta' \bmod p^\theta} \rho_\theta(\xi')^2\rho_\theta(\eta')^2 \oint_{p^B} |\mathfrak{f}_\theta(\boldsymbol{\alpha};\xi')^2 \mathfrak{f}_\theta(\boldsymbol{\alpha};\eta')^{2s-2}|\,\mathrm{d}\boldsymbol{\alpha},$$

where the summation over ξ' and η' is subject to the conditions

$$\xi' \equiv \xi \ (\mathrm{mod}\ p^\nu) \quad \text{and} \quad \eta' \equiv \eta \ (\mathrm{mod}\ p^\nu).$$

Thus $K_{\nu,\nu,c}^{1,\boldsymbol{\varphi},\nu}(\mathfrak{a};\xi,\eta)$ is at most

$$\rho_\nu(\xi)^{-2}\rho_\nu(\eta)^{-2}(p^{\theta-\nu})^s \sum_{\xi',\eta' \bmod p^\theta} \rho_\theta(\xi')^2\rho_\theta(\eta')^2 K_{\theta,\theta,c}^{1,\boldsymbol{\varphi},\nu}(\mathfrak{a};\xi',\eta').$$

In view of the definition (3.19), we find that

$$K_{\nu,\nu,c}^{1,\boldsymbol{\varphi},\nu}(\mathfrak{a}) \leqslant \rho_0^{-4}(p^{\theta-\nu})^s \sum_{\xi' \bmod p^\theta} \sum_{\substack{\eta' \bmod p^\theta \\ \eta' \not\equiv \xi' \ (\mathrm{mod}\ p^\nu)}} \rho_\theta(\xi')^2\rho_\theta(\eta')^2 K_{\theta,\theta,c}^{1,\boldsymbol{\varphi},\nu}(\mathfrak{a};\xi',\eta')$$
$$= (p^{\theta-\nu})^s K_{\theta,\theta,c}^{1,\boldsymbol{\varphi},\nu}(\mathfrak{a}).$$

By substituting this bound into the estimate supplied by Lemma 6.1, we arrive at the bound

$$U_{s,k}^B(\mathfrak{a}) \ll p^{s(\theta-\nu)+s\nu} K_{\theta,\theta,c}^{1,\boldsymbol{\varphi},\nu}(\mathfrak{a}),$$

and the conclusion of the lemma follows at once. □

7. Harnessing approximate translation-dilation invariance

This is the section of the paper that does the heavy lifting, for we now initiate our iterative process. Throughout, we suppose that $k \geqslant 2$, and we assume the validity of Theorem 3.1 for exponents smaller than k. Our strategy is to approximate the system of congruences underlying the mean value $K^{r,\boldsymbol{\varphi},\nu}_{a,b,c}(\mathfrak{a})$ by a corresponding system of Vinogradov type. As in our earlier efficient congruencing methods, we are able to extract a strong congruence condition on the underlying variables. Throughout, the conventions of the previous section remain in play, and in particular we write $R = r(r+1)/2$.

Lemma 7.1. *Suppose that a, b and r are integers with*

$$1 \leqslant r \leqslant k-1, \quad a \geqslant \delta\theta, \quad b \geqslant \delta\theta \quad \text{and} \quad ra \leqslant (k-r+1)b \leqslant B. \tag{7.1}$$

Put

$$b' = \lceil (k-r+1)b/r \rceil. \tag{7.2}$$

Then one has

$$K^{r,\boldsymbol{\varphi},\nu}_{a,b,c}(\mathfrak{a}) \ll p^{k^2\nu} K^{r,\boldsymbol{\varphi},\nu}_{b',b,c}(\mathfrak{a}).$$

Proof. In view of the definition (3.19), we focus initially on the mean value $K^{r,\boldsymbol{\varphi},\nu}_{a,b,c}(\mathfrak{a};\xi,\eta)$, in which we may suppose that $\eta \not\equiv \xi \pmod{p^\nu}$. The latter condition permits us the hypothesis $p^\gamma \| (\xi - \eta)$ for some non-negative integer γ with $\gamma < \nu$. We define the integer ω by putting $\omega = (\xi - \eta)p^{-\gamma}$, so that one has $(\omega, p) = 1$. It is convenient for future use to introduce the parameter

$$B' = (k-r+1)b - ra - (k-r)\gamma. \tag{7.3}$$

It transpires that in our main argument it is useful to have a little extra room in which to work. We therefore begin by exploring the situation in which $B' \leqslant \nu$. In such circumstances, it follows from (7.3) that

$$(k-r+1)b - ra \leqslant \nu + (k-r)\gamma \leqslant k\nu.$$

On recalling (7.2), one then finds that

$$b' - a \leqslant 1 + k\nu/r, \tag{7.4}$$

and an application of Lemma 6.2 leads to the bound

$$\rho_a(\xi)^2 |\mathfrak{f}_a(\boldsymbol{\alpha};\xi)|^{2R} \leqslant (p^{b'-a})^R \sum_{\substack{\zeta \bmod p^{b'} \\ \zeta \equiv \xi \ (\mathrm{mod}\ p^a)}} \rho_{b'}(\zeta)^2 |\mathfrak{f}_{b'}(\boldsymbol{\alpha};\zeta)|^{2R}.$$

Thus we discern from (3.19) and (3.20) that

$$K^{r,\boldsymbol{\varphi},\nu}_{a,b,c}(\mathfrak{a}) \leqslant (p^{b'-a})^R K^{r,\boldsymbol{\varphi},\nu}_{b',b,c}(\mathfrak{a}).$$

Consequently, in view of (7.4) and the relation $R = r(r+1)/2 < kr/2$, we obtain the estimate

$$K^{r,\boldsymbol{\varphi},\nu}_{a,b,c}(\mathfrak{a}) \leqslant p^{k^2\nu} K^{r,\boldsymbol{\varphi},\nu}_{b',b,c}(\mathfrak{a}).$$

This bound delivers the conclusion of the lemma when $B' \leqslant \nu$.

We now focus on the alternate situation with

$$B' > \nu. \tag{7.5}$$

We recall that, by orthogonality and the definition (3.20), the mean value $K_{a,b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a};\xi,\eta)$ counts the integral solutions of the simultaneous congruences (3.21), with their attendant conditions (3.22), and with each solution being counted with weight (3.23). We temporarily focus exclusively on the nature of these congruences.

In the first phase of our argument, we seek to extract from the right hand side of (3.21) a system resembling one of Vinogradov type. Observe that $\boldsymbol{\varphi}$ is a p^c-spaced system, so it follows just as in the argument of the proof of Lemma 4.1 leading to (4.6) that, in the first instance, there is no loss of generality in supposing that

$$\varphi_j(t) = t^j + p^c t^{k+1}\psi_j(t) \quad (1 \leqslant j \leqslant k),$$

for appropriate polynomials $\psi_j \in \mathbb{Z}[t]$. In view of the constraints (3.22), we may substitute

$$v_l = p^b u_l + \eta \quad \text{and} \quad w_l = p^b z_l + \eta \quad (1 \leqslant l \leqslant s - R),$$

and (3.21) is transformed into the new system

$$\sum_{i=1}^{R} (\varphi_j(x_i) - \varphi_j(y_i)) \equiv \sum_{l=1}^{s-R} \big(\varphi_j(p^b u_l + \eta) - \varphi_j(p^b z_l + \eta)\big) \pmod{p^B}$$
$$(1 \leqslant j \leqslant k). \quad (7.6)$$

Next, as in the argument of the proof of Lemma 4.1 leading from (4.6) to (4.7) and (4.8), we may take appropriate integral linear combinations of the congruences comprising (7.6) to confirm that there exist polynomials $\Psi_j \in \mathbb{Z}[t]$ for which

$$\sum_{i=1}^{R} (\Phi_j(x_i - \eta) - \Phi_j(y_i - \eta)) \equiv \sum_{l=1}^{s-R} \big(\Phi_j(p^b u_l) - \Phi_j(p^b z_l)\big) \pmod{p^B}$$
$$(1 \leqslant j \leqslant k),$$

in which

$$\Phi_j(t) = t^j + p^c t^{k+1}\Psi_j(t). \quad (7.7)$$

In particular, the system $\boldsymbol{\Phi}$ is p^c-spaced, and one has

$$\sum_{i=1}^{R} \Phi_j(x_i - \eta) \equiv \sum_{i=1}^{R} \Phi_j(y_i - \eta) \pmod{(p^{jb}, p^B)} \quad (1 \leqslant j \leqslant k). \quad (7.8)$$

In the second phase of our analysis, we seek to extract from (7.8) a system resembling one of Vinogradov type. With a limited supply of variable names available, it will be expedient to recycle letters in circumstances where confusion is easily avoided. We recall from (3.22) that in the solutions of (3.21) of interest to us, we have $\mathbf{x} \equiv \mathbf{y} \equiv \xi \pmod{p^a}$. Thus we may substitute

$$x_i = p^a u_i + \xi \quad \text{and} \quad y_i = p^a z_i + \xi \quad (1 \leqslant i \leqslant R). \quad (7.9)$$

We recall that $\xi - \eta = \omega p^\gamma$, and that we may suppose that $B \geqslant (k - r + 1)b$. Thus, by dropping mention in (7.8) of those congruences with index j satisfying

$1 \leqslant j \leqslant k-r$, and substituting (7.9) into (7.8), we deduce that

$$\sum_{i=1}^{R} \Phi_j(p^a u_i + \omega p^\gamma) \equiv \sum_{i=1}^{R} \Phi_j(p^a z_i + \omega p^\gamma) \pmod{p^{(k-r+1)b}}$$
$$(k-r+1 \leqslant j \leqslant k). \tag{7.10}$$

The situation in (7.10) is somewhat akin to that encountered in the argument leading from (4.6) to (4.7) and (4.8). We apply the binomial theorem within (7.7). In this way, when $1 \leqslant l \leqslant r$, we find that for suitable polynomials $\Theta_l \in \mathbb{Z}[t]$, one has

$$\Phi_{k-r+l}(p^a y + \omega p^\gamma) - \Phi_{k-r+l}(\omega p^\gamma) = \Upsilon_l(p^a y),$$

in which

$$\Upsilon_l(t) = \sum_{i=1}^{r} \Omega_{il} (\omega p^\gamma)^{k-r+l-i} t^i + t^{r+1} \Theta_l(t),$$

and the integral coefficients Ω_{il} satisfy the congruence

$$\Omega_{il} \equiv \binom{k-r+l}{i} \pmod{p^c} \quad (1 \leqslant i, l \leqslant r).$$

Much as in the argument of the proof of [49, Lemma 3.2], our hypothesis $p > k$ ensures that the matrix $A = (\Omega_{il})_{1 \leqslant i,l \leqslant r}$ satisfies $\det(A) \not\equiv 0 \pmod{p}$, and hence that A possesses a multiplicative inverse A^{-1} modulo $p^{(k-r+1)b}$ having integral coefficients. Indeed, it is apparent that $(1! \cdot 2! \cdots r!) \det(A)$ is congruent modulo p to

$$\det\left((k-j+1) \cdots (k-j+1-(i-1))\right)_{1 \leqslant i,j \leqslant r} = \det\left((k-j+1)^i\right)_{1 \leqslant i,j \leqslant r},$$

and hence also to

$$\left(\prod_{l=1}^{r} (k-l+1)\right)\left(\prod_{1 \leqslant i < j \leqslant r} ((k-j+1) - (k-i+1))\right) \not\equiv 0 \pmod{p}.$$

Here, we have corrected an inconsequential oversight in the evaluation of a determinant in the proof of [49, Lemma 3.2].

We now replace $\boldsymbol{\Upsilon}$ by $A^{-1}\boldsymbol{\Upsilon}$ and $\boldsymbol{\Theta}$ by $A^{-1}\boldsymbol{\Theta}$. This once again amounts to taking suitable integral linear combinations of the congruences comprising (7.10). In this way, we see that there is no loss of generality in supposing that the coefficient matrix A is equal to I_r. Since $(\omega, p) = 1$, moreover, there is an integral multiplicative inverse ω^{-1} for ω modulo $p^{(k-r+1)b}$. Hence, there exist polynomials $\Xi_l \in \mathbb{Z}[t]$ having the property that whenever the system (7.10) is satisfied, then

$$(\omega p^\gamma)^{k-r} \sum_{i=1}^{R} (p^a)^l \left(\Psi_l(u_i) - \Psi_l(z_i)\right) \equiv 0 \pmod{p^{(k-r+1)b}} \quad (1 \leqslant l \leqslant r),$$

in which

$$\Psi_l(t) = t^l + p^{a-(k-r)\gamma} \Xi_l(t). \tag{7.11}$$

Notice here that, since we may suppose from (7.1) that $a \geqslant \delta\theta$, and $\gamma \leqslant \nu$, our hypothesis concerning the hierarchy (6.2) combines with (6.5) and (6.10) to give

$$k\gamma \leqslant k\lceil 4\varepsilon H\Lambda^{-1}\rceil < \delta^2\lceil \mu H\rceil \leqslant \delta a,$$

and hence

$$a-(k-r)\gamma > (1-\delta)a \geqslant \delta(1-\delta)\mu H > \tau B.$$

The exponent of p on the right hand side of (7.11) is therefore positive, and indeed the system $\boldsymbol{\Psi}$ is p^c-spaced for some $c > \tau(k-r+1)b$.

The integral solutions $\mathbf{x} = p^a\mathbf{u}+\xi$, $\mathbf{y} = p^a\mathbf{z}+\xi$, $\mathbf{v}$, $\mathbf{w}$ of the original system of congruences (3.21), counted with weight (3.23) associated with the definition (3.20) of $K^{r,\boldsymbol{\varphi},\nu}_{a,b,c}(\mathfrak{a};\xi,\eta)$, are therefore constrained by the additional system of congruences

$$\sum_{i=1}^{R}\Psi_l(u_i) \equiv \sum_{i=1}^{R}\Psi_l(z_i) \pmod{p^{B'}} \quad (1 \leqslant l \leqslant r), \tag{7.12}$$

in which B' is defined by (7.3).

Our goal at this point is to apply the inductive hypothesis, meaning the conclusion of Theorem 3.1 with k replaced by r, so as to extract congruence information from the relations (7.12). We recall at this point that in view of (7.5), we may suppose that $B' > \nu$. Write $H' = \lceil B'/r\rceil$. Then in view of (7.2) and (7.3), one has

$$b' - H' \leqslant a+1+(k-r)\gamma/r. \tag{7.13}$$

We note from (6.5) that $B' > 4\varepsilon H\Lambda^{-1}$, so our hypotheses concerning B and the hierarchy (6.2) ensure that B' is sufficiently large in terms of k, Λ, μ, τ, δ and ε. We reinterpret the definition (3.4) of $\mathfrak{f}_h(\boldsymbol{\alpha};\xi)$ in the shape (4.2), so that

$$\mathfrak{f}_a(\boldsymbol{\alpha};\xi) = \rho_a(\xi)^{-1}\sum_{y\in\mathbb{Z}}\mathfrak{c}_y(\boldsymbol{\alpha}), \tag{7.14}$$

where

$$\mathfrak{c}_y(\boldsymbol{\alpha}) = \mathfrak{a}_{p^ay+\xi}e(\psi(p^ay+\xi;\boldsymbol{\alpha})) \quad (y\in\mathbb{Z}),$$

and $\psi(n;\boldsymbol{\alpha})$ is defined by (3.3). Notice that

$$\rho_0(1;\mathfrak{c})^2 = \sum_{y\in\mathbb{Z}}|\mathfrak{c}_y(\boldsymbol{\alpha})|^2 = \sum_{y\in\mathbb{Z}}|\mathfrak{a}_{p^ay+\xi}|^2 = \rho_a(\xi)^2. \tag{7.15}$$

Finally, define the exponential sum

$$\mathfrak{g}_{\mathfrak{c}}(\boldsymbol{\alpha},\boldsymbol{\beta}) = \rho_0(1;\mathfrak{c})^{-1}\sum_{y\in\mathbb{Z}}\mathfrak{c}_y(\boldsymbol{\alpha})e(\beta_1\Psi_1(y)+\ldots+\beta_r\Psi_r(y)), \tag{7.16}$$

and the mean value

$$J(\boldsymbol{\alpha}) = \oint_{[0,1)^{B'}}|\mathfrak{g}_{\mathfrak{c}}(\boldsymbol{\alpha},\boldsymbol{\beta})|^{2R}\,\mathrm{d}\boldsymbol{\beta}.$$

By orthogonality, the mean value $J(\boldsymbol{\alpha})$ counts the integral solutions of the system of congruences (7.12), with each solution $\mathbf{u}$, $\mathbf{z}$ being counted with weight

$$\begin{aligned}&\rho_0(1;\mathfrak{c})^{-2R}\prod_{i=1}^{R}\mathfrak{c}_{u_i}(\boldsymbol{\alpha})\overline{\mathfrak{c}}_{z_i}(\boldsymbol{\alpha})\\&=\rho_a(\xi)^{-2R}\Big(\prod_{i=1}^{R}\mathfrak{a}_{p^au_i+\xi}\overline{\mathfrak{a}}_{p^az_i+\xi}\Big)e\Big(\sum_{i=1}^{R}(\psi(p^au_i+\xi;\boldsymbol{\alpha})-\psi(p^az_i+\xi;\boldsymbol{\alpha}))\Big).\end{aligned}$$

But the conditions (7.12) on $\mathbf{x}=p^a\mathbf{u}+\xi$ and $\mathbf{y}=p^a\mathbf{z}+\xi$ are implied by (3.21), as we have seen. Thus, on noting that from (7.14)-(7.16), one has

$$\mathfrak{g}_{\mathfrak{c}}(\boldsymbol{\alpha},\mathbf{0})=\mathfrak{f}_a(\boldsymbol{\alpha};\xi),$$

we deduce that

$$\oint_{p^B}|\mathfrak{f}_a(\boldsymbol{\alpha};\xi)^{2R}\mathfrak{f}_b(\boldsymbol{\alpha};\eta)^{2s-2R}|\,\mathrm{d}\boldsymbol{\alpha}=\oint_{p^B}\oint_{p^{B'}}|\mathfrak{g}_{\mathfrak{c}}(\boldsymbol{\alpha},\boldsymbol{\beta})^{2R}\mathfrak{f}_b(\boldsymbol{\alpha};\eta)^{2s-2R}|\,\mathrm{d}\boldsymbol{\beta}\,\mathrm{d}\boldsymbol{\alpha},\tag{7.17}$$

whence

$$K^{r,\boldsymbol{\varphi},\nu}_{a,b,c}(\mathfrak{a};\xi,\eta)=\oint_{p^B}J(\boldsymbol{\alpha})|\mathfrak{f}_b(\boldsymbol{\alpha};\eta)|^{2s-2R}\,\mathrm{d}\boldsymbol{\alpha}.\tag{7.18}$$

The point here is that the inner integral over $\boldsymbol{\beta}$ on the right hand side of (7.17) is essentially redundant, since the integral over $\boldsymbol{\alpha}$ already restricts the variables underlying the term $|\mathfrak{g}_{\mathfrak{c}}(\boldsymbol{\alpha},\boldsymbol{\beta})|^{2R}$ to satisfy the system of congruences (7.12).

The mean value $J(\boldsymbol{\alpha})$ is of the type estimated in Corollary 3.2 to Theorem 3.1. The system $\boldsymbol{\Psi}$ is p^c-spaced for some $c\geqslant\tau B'$ and B' is sufficiently large in terms of k, τ and ε^2. Hence

$$\begin{aligned}J(\boldsymbol{\alpha})=U^{B'}_{R,r}(\mathfrak{c})&\ll p^{B'\varepsilon^2}U^{B',H'}_{R,r}(\mathfrak{c})\\&=p^{B'\varepsilon^2}\rho_0(1;\mathfrak{c})^{-2}\sum_{\zeta\bmod p^{H'}}\rho_{H'}(\zeta;\mathfrak{c})^2\oint_{p^{B'}}|\mathfrak{g}_{\mathfrak{c}'}(\boldsymbol{\alpha},\boldsymbol{\beta})|^{2R}\,\mathrm{d}\boldsymbol{\beta},\end{aligned}\tag{7.19}$$

in which the sequence $\mathfrak{c}'=(\mathfrak{c}'_n(\boldsymbol{\alpha}))_{n\in\mathbb{Z}}$ is defined by putting

$$\mathfrak{c}'_n(\boldsymbol{\alpha})=\begin{cases}\mathfrak{c}_n(\boldsymbol{\alpha}), & \text{when } n\equiv\zeta\ (\mathrm{mod}\ p^{H'}),\\ 0, & \text{otherwise.}\end{cases}$$

We have $\rho_0(1;\mathfrak{c})=\rho_a(\xi)$. Also, on writing $\kappa=p^a\zeta+\xi$, we see that

$$\rho_{H'}(\zeta;\mathfrak{c})^2=\sum_{y\equiv\zeta\ (\mathrm{mod}\ p^{H'})}|\mathfrak{a}_{p^ay+\xi}|^2=\sum_{n\equiv\kappa\ (\mathrm{mod}\ p^{a+H'})}|\mathfrak{a}_n|^2=\rho_{a+H'}(\kappa)^2.$$

Thus, on substituting (7.19) into (7.18), we conclude that

$$\begin{aligned}&\rho_a(\xi)^2K_{a,b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a};\xi,\eta)\\&\ll p^{B'\varepsilon^2}\sum_{\substack{\kappa \bmod p^{a+H'}\\ \kappa\equiv\xi\ (\mathrm{mod}\ p^a)}}\rho_{a+H'}(\kappa)^2\oint_{p^B}\oint_{p^{B'}}|\mathfrak{g}_{\mathfrak{c}'}(\boldsymbol{\alpha},\boldsymbol{\beta})^{2R}\mathfrak{f}_b(\boldsymbol{\alpha};\eta)^{2s-2R}|\,\mathrm{d}\boldsymbol{\beta}\,\mathrm{d}\boldsymbol{\alpha}.\end{aligned}\tag{7.20}$$

We next unravel the mean value occurring on the right hand side of (7.20), reversing our previous course. The mean value

$$\oint_{p^{B'}}|\mathfrak{g}_{\mathfrak{c}'}(\boldsymbol{\alpha},\boldsymbol{\beta})|^{2R}\,\mathrm{d}\boldsymbol{\beta}$$

counts the integral solutions of the system of congruences (7.12), with each solution $\mathbf{u}$, $\mathbf{z}$ being counted with weight

$$\begin{aligned}&\rho_0(1;\mathfrak{c}')^{-2R}\prod_{i=1}^R\mathfrak{c}'_{u_i}(\boldsymbol{\alpha})\overline{\mathfrak{c}}'_{z_i}(\boldsymbol{\alpha})\\&=\rho_{a+H'}(\kappa)^{-2R}\Big(\prod_{i=1}^R\mathfrak{a}_{p^au_i+\xi}\overline{\mathfrak{a}}_{p^az_i+\xi}\Big)e\Big(\sum_{i=1}^R(\psi(p^au_i+\xi;\boldsymbol{\alpha})-\psi(p^az_i+\xi;\boldsymbol{\alpha}))\Big),\end{aligned}$$

but now subject to the constraint $\mathbf{u}\equiv\mathbf{z}\equiv\zeta\ (\mathrm{mod}\ p^{H'})$. The conditions (7.12) on $\mathbf{x}=p^a\mathbf{u}+\xi$ and $\mathbf{y}=p^a\mathbf{z}+\xi$ are again implied by (3.21). Then since

$$\mathfrak{g}_{\mathfrak{c}'}(\boldsymbol{\alpha};\mathbf{0})=\mathfrak{f}_{a+H'}(\boldsymbol{\alpha};\kappa),$$

we deduce that

$$\begin{aligned}&\oint_{p^B}\oint_{p^{B'}}|\mathfrak{g}_{\mathfrak{c}'}(\boldsymbol{\alpha},\boldsymbol{\beta})^{2R}\mathfrak{f}_b(\boldsymbol{\alpha};\eta)^{2s-2R}|\,\mathrm{d}\boldsymbol{\beta}\,\mathrm{d}\boldsymbol{\alpha}\\&\qquad=\oint_{p^B}|\mathfrak{f}_{a+H'}(\boldsymbol{\alpha};\kappa)^{2R}\mathfrak{f}_b(\boldsymbol{\alpha};\eta)^{2s-2R}|\,\mathrm{d}\boldsymbol{\alpha}.\end{aligned}\tag{7.21}$$

Moreover, on making use of Lemma 6.2, one sees that

$$\rho_{a+H'}(\kappa)^2|\mathfrak{f}_{a+H'}(\boldsymbol{\alpha};\kappa)|^{2R}\leqslant\Big(p^{b'-(a+H')}\Big)^R\sum_{\substack{\xi' \bmod p^{b'}\\ \xi'\equiv\kappa\ (\mathrm{mod}\ p^{a+H'})}}\rho_{b'}(\xi')^2|\mathfrak{f}_{b'}(\boldsymbol{\alpha};\xi')|^{2R}.\tag{7.22}$$

Thus, on substituting (7.22) into (7.21) and thence into (7.20), we obtain the bound

$$\rho_a(\xi)^2K_{a,b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a};\xi,\eta)\ll p^{B'\varepsilon^2+R(b'-a-H')}\sum_{\substack{\xi' \bmod p^{b'}\\ \xi'\equiv\xi\ (\mathrm{mod}\ p^a)}}\rho_{b'}(\xi')^2K_{b',b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a};\xi',\eta).\tag{7.23}$$

Next, substituting (7.23) into (3.19), we conclude that

$$\rho_0^4 K_{a,b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a}) \ll p^{B'\varepsilon^2+R(b'-a-H')} \sum_{\xi' \bmod p^{b'}} \sum_{\substack{\eta \bmod p^b \\ \eta \not\equiv \xi' \pmod{p^\nu}}} \rho_{b'}(\xi')^2 \rho_b(\eta)^2 K_{b',b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a};\xi',\eta),$$

whence

$$K_{a,b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a}) \ll p^{B'\varepsilon^2+R(b'-a-H')} K_{b',b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a}). \tag{7.24}$$

In order to complete the proof of the lemma, it remains only to note that, from (7.13), one has

$$\begin{aligned} R(b'-a-H') &= r(r+1)(b'-a-H')/2 \\ &\leqslant (r+1)(r+(k-r)\gamma)/2 < k^2\nu/2. \end{aligned}$$

Thus, in view of the definition (6.5) and our hierarchy (6.2),

$$B'\varepsilon^2 + R(b'-a-H') < \nu + k^2\nu/2 < k^2\nu.$$

The conclusion of the lemma therefore follows from (7.24) in this second situation with $B' > \nu$. □

8. The iterative step

The work of the previous section shows how to generate powerful congruence constraints on the variables underlying the mean value $K_{a,b,c}^{r,\boldsymbol{\varphi},\nu}(\mathfrak{a})$. As in earlier efficient congruencing methods, we must now interchange the roles of the two sets of variables so that these congruence constraints may be employed anew to generate yet stronger constraints. In our earlier multigrade method, this was achieved by employing Hölder's inequality in a somewhat greedy manner. On this occasion, we take a slightly more measured approach, though the underlying ideas remain the same. At this point, the parameters $\boldsymbol{\varphi}$, ν and c have ceased to possess any particular significance, and we omit mention of them in our various notations. Also, throughout this section, we suppose that a, b and r are integers satisfying (7.1), and we define b' via (7.2).

Lemma 8.1. *When $r \geqslant 2$, one has*

$$K_{a,b}^r(\mathfrak{a}) \ll p^{k^2\nu} K_{b,b'}^{k-r}(\mathfrak{a})^{\frac{1}{k-r+1}} K_{b',b}^{r-1}(\mathfrak{a})^{\frac{k-r}{k-r+1}}. \tag{8.1}$$

Meanwhile, when $r = 1$, one instead has

$$K_{a,b}^1(\mathfrak{a}) \ll p^{k^2\nu} K_{b,kb}^{k-1}(\mathfrak{a})^{1/k} U_{s,k}^{B,b}(\mathfrak{a})^{1-1/k}. \tag{8.2}$$

Proof. Observe that

$$\begin{aligned} k(k+1) - r(r+1) &= (k+r)(k-r) + (k-r) \\ &= \frac{(k-r)(k-r+1)}{k-r+1} + (k(k+1)-r(r-1)) \cdot \frac{k-r}{k-r+1}. \end{aligned}$$

Then for any integers ζ and η, it is a consequence of Hölder's inequality that when $1 \leqslant r \leqslant k-1$, one has

$$\oint_{p^B} |\mathfrak{f}_{b'}(\boldsymbol{\alpha};\zeta)^{r(r+1)} \mathfrak{f}_b(\boldsymbol{\alpha};\eta)^{k(k+1)-r(r+1)}| \, \mathrm{d}\boldsymbol{\alpha} \leqslant U_1^{\frac{1}{k-r+1}} U_2^{\frac{k-r}{k-r+1}},$$

where

$$U_1 = \oint_{p^B} |\mathfrak{f}_b(\boldsymbol{\alpha};\eta)^{(k-r)(k-r+1)}\mathfrak{f}_{b'}(\boldsymbol{\alpha};\zeta)^{k(k+1)-(k-r)(k-r+1)}|\,\mathrm{d}\boldsymbol{\alpha}$$

and

$$U_2 = \oint_{p^B} |\mathfrak{f}_{b'}(\boldsymbol{\alpha};\zeta)^{r(r-1)}\mathfrak{f}_b(\boldsymbol{\alpha};\eta)^{k(k+1)-r(r-1)}|\,\mathrm{d}\boldsymbol{\alpha}.$$

Thus, on recalling (3.19) and (3.20), we deduce first that

$$K^r_{b',b}(\mathfrak{a};\zeta,\eta) \leqslant K^{k-r}_{b,b'}(\mathfrak{a};\eta,\zeta)^{\frac{1}{k-r+1}} K^{r-1}_{b',b}(\mathfrak{a};\zeta,\eta)^{\frac{k-r}{k-r+1}},$$

and hence, by another application of Hölder's inequality, that

$$K^r_{b',b}(\mathfrak{a}) \leqslant K^{k-r}_{b,b'}(\mathfrak{a})^{\frac{1}{k-r+1}} K^{r-1}_{b',b}(\mathfrak{a})^{\frac{k-r}{k-r+1}}. \tag{8.3}$$

Since Lemma 7.1 shows that $K^r_{a,b}(\mathfrak{a}) \ll p^{k^2\nu}K^r_{b',b}(\mathfrak{a})$, the conclusion (8.1) of the lemma is immediate from (8.3) in the case $r \geqslant 2$.

In order to handle the case $r = 1$, we begin by noting that from (3.19) and (3.20), one has

$$K^0_{b',b}(\mathfrak{a}) = \rho_0^{-4} \sum_{\xi \bmod p^{b'}} \sum_{\substack{\eta \bmod p^b \\ \xi \not\equiv \eta \pmod{p^\nu}}} \rho_{b'}(\xi)^2\rho_b(\eta)^2 \oint_{p^B} |\mathfrak{f}_b(\boldsymbol{\alpha};\eta)|^{2s}\,\mathrm{d}\boldsymbol{\alpha}.$$

Since it follows from (3.8) that

$$\sum_{\eta \bmod p^b} \rho_b(\eta)^2 \oint_{p^B} |\mathfrak{f}_b(\boldsymbol{\alpha};\eta)|^{2s}\,\mathrm{d}\boldsymbol{\alpha} = \rho_0^2 U^{B,b}_{s,k}(\mathfrak{a})$$

and

$$\sum_{\xi \bmod p^{b'}} \rho_{b'}(\xi)^2 = \rho_0^2,$$

it follows that $K^0_{b',b}(\mathfrak{a}) = U^{B,b}_{s,k}(\mathfrak{a})$, and so the desired conclusion (8.2) again follows from (8.3) when $r = 1$. □

We next interpret the conclusion of Lemma 8.1 in terms of the anticipated order of magnitude normalisation defined in (3.24).

Lemma 8.2. *Suppose that $b \leqslant (1-\delta)B/k$. Then, when $r \geqslant 2$, one has*

$$[\![K^r_{a,b}(\mathfrak{a})]\!]_\Lambda \ll p^{k^2\nu}[\![K^{k-r}_{b,b'}(\mathfrak{a})]\!]_\Lambda^{1/(k-r+1)}[\![K^{r-1}_{b',b}(\mathfrak{a})]\!]_\Lambda^{1-1/r}. \tag{8.4}$$

Meanwhile, when $r = 1$ one instead has

$$[\![K^1_{a,b}(\mathfrak{a})]\!]_\Lambda \ll p^{2k^2\nu}[\![K^{k-1}_{b,kb}(\mathfrak{a})]\!]_\Lambda^{1/k}(p^{-b})^{\Lambda(1-1/k)}. \tag{8.5}$$

Proof. By reference to (3.24), we deduce from Lemma 8.1 that for $r \geqslant 2$, one has

$$[\![K^r_{a,b}(\mathfrak{a})]\!]_\Lambda \ll (p^{k^2\nu})^{\frac{k-1}{r(k-r)}} V_1^{\frac{1}{k-r+1}} V_2^{\frac{k-r}{k-r+1}},$$

where

$$V_1 = \left(\frac{K_{b,b'}^{k-r}(\mathfrak{a})}{p^{\Lambda H} U_{s,k}^{B,H}(\mathfrak{a})} \right)^{\frac{k-1}{r(k-r)}} = [\![K_{b,b'}^{k-r}(\mathfrak{a})]\!]_\Lambda,$$

and

$$V_2 = \left(\frac{K_{b',b}^{r-1}(\mathfrak{a})}{p^{\Lambda H} U_{s,k}^{B,H}(\mathfrak{a})} \right)^{\frac{k-1}{r(k-r)}} = [\![K_{b',b}^{r-1}(\mathfrak{a})]\!]_\Lambda^{\binom{r-1}{r}\binom{k-r+1}{k-r}}.$$

The first conclusion (8.4) is now immediate for $r \geqslant 2$. Here, we have made use of the elementary fact that

$$\max_{1 \leqslant r \leqslant k-1} \frac{k-1}{r(k-r)} = 1.$$

When $r = 1$, on the other hand, one finds in like manner that

$$[\![K_{a,b}^1(\mathfrak{a})]\!]_\Lambda \ll p^{k^2\nu} V_3^{1/k} V_4^{1-1/k}, \tag{8.6}$$

where

$$V_3 = \left(\frac{K_{b,kb}^{k-1}(\mathfrak{a})}{p^{\Lambda H} U_{s,k}^{B,H}(\mathfrak{a})} \right) = [\![K_{b,kb}^{k-1}(\mathfrak{a})]\!]_\Lambda \tag{8.7}$$

and

$$V_4 = \frac{U_{s,k}^{B,b}(\mathfrak{a})}{p^{\Lambda H} U_{s,k}^{B,H}(\mathfrak{a})}.$$

In view of the upper bound (6.4), one has

$$U_{s,k}^{B,b}(\mathfrak{a}) \leqslant (p^{H-b})^{\Lambda+\varepsilon} U_{s,k}^{B,H}(\mathfrak{a}),$$

and hence, on recalling (6.1) and (6.5), one finds that

$$V_4 \leqslant p^{(H-b)\varepsilon - \Lambda b} \leqslant p^{\Lambda\nu - \Lambda b} \leqslant p^{s\nu - \Lambda b}.$$

On substituting this bound together with (8.7) into (8.6), we conclude that

$$[\![K_{a,b}^1(\mathfrak{a})]\!]_\Lambda \ll p^{(k^2+s)\nu} [\![K_{b,kb}^{k-1}(\mathfrak{a})]\!]_\Lambda^{1/k} (p^{-b})^{\Lambda(1-1/k)},$$

and the conclusion (8.5) of the lemma follows on recalling that

$$s = k(k+1)/2 \leqslant k^2.$$

□

9. Distilling multigrade combinations into monograde processes

We turn our attention next to the problem of analysing the impact of applying Lemma 8.2 iteratively so as to estimate $K_{a,b}^r(\mathfrak{a})$ in terms of a tree of possible outcomes. Such processes seem, at first appearance, difficult to control. However, the multigrade efficient congruencing method of [54, 55, 58] offers the tools to accommodate such an analysis. In broad terms, we weight the possible outcomes in such a manner that one can follow any single path through the tree, and compute the outcome without reference to the multitude of alternate paths available.

In this section and the next, in the interests of concision, we write

$$\widetilde{K}^r_{a,b} = [\![K^r_{a,b}(\mathfrak{a})]\!]_\Lambda.$$

Also, when $1 \leqslant j \leqslant k-1$, we write

$$\rho_j = \frac{j}{k-j+1} \quad \text{and} \quad b_j = \left\lceil \frac{b}{\rho_j} \right\rceil.$$

Lemma 9.1. *Suppose that*

$$1 \leqslant r \leqslant k-1, \quad a \geqslant \delta\theta, \quad b \geqslant k\delta\theta \quad \textit{and} \quad ra \leqslant (k-r+1)b. \tag{9.1}$$

Then whenever $kb \leqslant (1-\delta)B$, one has

$$\widetilde{K}^r_{a,b} \ll p^{(r+1)k^2\nu}(p^{-b})^{(1-1/k)\Lambda/r} \prod_{j=1}^{r} \left(\widetilde{K}^{k-j}_{b,b_j}\right)^{\rho_j/r}. \tag{9.2}$$

Proof. We establish (9.2) by induction on r. Observe that in the case $r = 1$, it follows from Lemma 8.2 that

$$\widetilde{K}^1_{a,b} \ll p^{2k^2\nu}(\widetilde{K}^{k-1}_{b,kb})^{1/k}(p^{-b})^{(1-1/k)\Lambda},$$

and this delivers (9.2) in this initial case. Note here that the hypotheses (9.1) together with the assumption $kb \leqslant (1-\delta)B$ imply the corresponding hypotheses (7.1) implicitly assumed in the statement of Lemma 8.2.

Suppose next that when $kb \leqslant (1-\delta)B$, the estimate (9.2) has been confirmed for all indices $r < \mathrm{P}$, for some integer P with $2 \leqslant \mathrm{P} \leqslant k-1$. In these circumstances, it follows from Lemma 8.2 that

$$\widetilde{K}^{\mathrm{P}}_{a,b} \ll p^{k^2\nu}(\widetilde{K}^{k-\mathrm{P}}_{b,b_{\mathrm{P}}})^{\rho_{\mathrm{P}}/\mathrm{P}}(\widetilde{K}^{\mathrm{P}-1}_{b_{\mathrm{P}},b})^{1-1/\mathrm{P}}, \tag{9.3}$$

with

$$b_{\mathrm{P}} = \left\lceil \frac{k-\mathrm{P}+1}{\mathrm{P}} b \right\rceil \geqslant \frac{2b}{k} > \delta\theta.$$

The latter lower bound, together with the upper bound

$$(\mathrm{P}-1)b_{\mathrm{P}} \leqslant (\mathrm{P}-1)\left(\frac{k-\mathrm{P}+1}{\mathrm{P}}b + 1\right) < (k-\mathrm{P}+2)b,$$

ensures that the conditions are satisfied permitting the inductive hypothesis (9.2) to be deployed to estimate $\widetilde{K}^{\mathrm{P}-1}_{b_{\mathrm{P}},b}$. Thus, we have

$$\widetilde{K}^{\mathrm{P}-1}_{b_{\mathrm{P}},b} \ll p^{\mathrm{P}k^2\nu}(p^{-b})^{(1-1/k)\Lambda/(\mathrm{P}-1)} \prod_{j=1}^{\mathrm{P}-1}(\widetilde{K}^{k-j}_{b,b_j})^{\rho_j/(\mathrm{P}-1)}.$$

On substituting this bound into (9.3), we deduce that

$$\widetilde{K}^{\mathrm{P}}_{a,b} \ll p^{\mathrm{P}k^2\nu}(p^{-b})^{(1-1/k)\Lambda/\mathrm{P}}(\widetilde{K}^{k-\mathrm{P}}_{b,b_{\mathrm{P}}})^{\rho_{\mathrm{P}}/\mathrm{P}} \prod_{j=1}^{\mathrm{P}-1}(\widetilde{K}^{k-j}_{b,b_j})^{\rho_j/\mathrm{P}},$$

and the inductive hypothesis (9.2) follows when $r = \mathrm{P}$. The desired conclusion (9.2) therefore follows by induction for $1 \leqslant r \leqslant k-1$. □

An immediate consequence of Lemma 9.1 provides an intermediate monograde iterative relation.

Lemma 9.2. *Suppose that* $1 \leqslant r \leqslant k-1$ *and* $kb \leqslant (1-\delta)B$. *Suppose further that the conditions (9.1) hold. Then there exists an integer* r' *with* $1 \leqslant r' \leqslant r$ *having the property that*

$$\widetilde{K}_{a,b}^{r} \ll (\widetilde{K}_{b,b_{r'}}^{k-r'})^{\rho_{r'}} (p^{-b})^{\Lambda/(2k)}.$$

Proof. By employing the elementary inequality

$$|z_1 \cdots z_n| \leqslant |z_1|^n + \ldots + |z_n|^n,$$

it follows from the relation (9.2) of Lemma 9.1 that

$$\widetilde{K}_{a,b}^{r} \ll p^{(r+1)k^2\nu} (p^{-b})^{(1-1/k)\Lambda/r} \sum_{j=1}^{r} (\widetilde{K}_{b,b_j}^{k-j})^{\rho_j}.$$

Thus we deduce that there exists an integer r' with $1 \leqslant r' \leqslant r$ such that

$$\widetilde{K}_{a,b}^{r} \ll p^{(r+1)k^2\nu} (p^{-b})^{(1-1/k)\Lambda/r} (\widetilde{K}_{b,b_{r'}}^{k-r'})^{\rho_{r'}}. \tag{9.4}$$

Note that when $1 \leqslant r \leqslant k-1$, one has $(1-1/k)/r \geqslant 1/k$. Moreover, in view of (6.5), (6.10), (9.1) and the hierarchy (6.2), one has

$$b\Lambda/k \geqslant \delta\theta\Lambda \geqslant \delta\mu H\Lambda > 2k^3\nu \geqslant 2(r+1)k^2\nu.$$

Thus we infer that

$$p^{(r+1)k^2\nu} (p^{-b})^{(1-1/k)\Lambda/r} \leqslant (p^{-b})^{\Lambda/(2k)}.$$

The conclusion of Lemma 9.2 follows by substituting this bound into (9.4). □

The estimate supplied by Lemma 9.2 bounds $\widetilde{K}_{a,b}^{r}$ in terms of $\widetilde{K}_{b,b_{r'}}^{k-r'}$, for a suitable integer r' with $1 \leqslant r' \leqslant r$. It is possible that the parameter $b_{r'}$ might be substantially smaller than b, and thus one might fear that continued iteration of such a relation might be limited by this shrinking parameter. At this point, we take the opportunity to dispell such fears once and for all.

Lemma 9.3. *Suppose that*

$$1 \leqslant r \leqslant k-1, \quad a \geqslant \delta\theta, \quad b \geqslant k^2\delta\theta \quad \text{and} \quad ra \leqslant (k-r+1)b. \tag{9.5}$$

Then, whenever $k^2 b \leqslant (1-\delta)B$, *there exist integers* a', b', r', *and there exists a positive number* ρ, *having the property that*

$$\widetilde{K}_{a,b}^{r} \ll (\widetilde{K}_{a',b'}^{r'})^{\rho} p^{-b\Lambda/(2k)}, \tag{9.6}$$

with

$$a' \geqslant \delta\theta, \quad b' \geqslant k^2\delta\theta, \quad r'a' \leqslant (k-r'+1)b', \tag{9.7}$$

$$1 \leqslant r' \leqslant k-1, \quad 0 < \rho < (1-1/k)^2, \tag{9.8}$$

$$(1+2/k)b \leqslant b' \leqslant k^2 b, \quad b' = \left\lceil \frac{r'+1}{k-r'} a' \right\rceil, \quad \rho b' \geqslant b. \tag{9.9}$$

Proof. We recall that we are permitted the assumption that a, b, r satisfy (9.5). Thus, as a consequence of Lemma 9.2, there exists an integer r_1 with $1 \leqslant r_1 \leqslant r$ having the property that

$$\widetilde{K}^r_{a,b} \ll (\widetilde{K}^{k-r_1}_{b,b_{r_1}})^{\rho_{r_1}}(p^{-b})^{\Lambda/(2k)}, \tag{9.10}$$

where

$$b_{r_1} = \left\lceil \frac{k-r_1+1}{r_1} b \right\rceil \quad \text{and} \quad \rho_{r_1} = \frac{r_1}{k-r_1+1}.$$

Notice here that $b \geqslant k^2\delta\theta \geqslant \delta\theta$,

$$b_{r_1} \geqslant b/k \geqslant k\delta\theta \quad \text{and} \quad b_{r_1} \leqslant kb \leqslant k^{-1}(1-\delta)B.$$

Moreover, one has

$$(k-r_1)b \leqslant (r_1+1)\left\lceil \frac{k-r_1+1}{r_1} b \right\rceil = (r_1+1)b_{r_1}.$$

We are therefore at liberty to apply Lemma 9.2 to estimate $\widetilde{K}^{k-r_1}_{b,b_{r_1}}$. Thus, there exists an integer r_2 with $1 \leqslant r_2 \leqslant k-r_1$ having the property that

$$\widetilde{K}^{k-r_1}_{b,b_{r_1}} \ll (\widetilde{K}^{k-r_2}_{b_{r_1},b_{r_2}})^{\rho_{r_2}}(p^{-b_{r_1}})^{\Lambda/(2k)}, \tag{9.11}$$

where

$$b_{r_2} = \left\lceil \frac{k-r_2+1}{r_2} b_{r_1} \right\rceil \quad \text{and} \quad \rho_{r_2} = \frac{r_2}{k-r_2+1}.$$

On substituting (9.11) into (9.10), we find that

$$\widetilde{K}^r_{a,b} \ll (\widetilde{K}^{r'}_{a',b'})^{\rho} p^{-\sigma\Lambda/(2k)}, \tag{9.12}$$

where

$$a' = b_{r_1}, \quad b' = b_{r_2}, \quad r' = k-r_2, \quad \rho = \rho_{r_1}\rho_{r_2} \quad \text{and} \quad \sigma = b + \rho_{r_1} b_{r_1}.$$

Since $\sigma \geqslant b$, the upper bound (9.12) will deliver the conclusion of the lemma provided that we are able to verify the conditions (9.7)-(9.9), a matter that we now address.

Observe first that since $1 \leqslant r_2 \leqslant k-r_1$ and $1 \leqslant r_1 \leqslant k-1$, one has

$$1 \leqslant r_1 \leqslant r' = k-r_2 \leqslant k-1.$$

Moreover,

$$\rho = \rho_{r_1}\rho_{r_2} = \frac{r_1}{k-r_1+1} \cdot \frac{r_2}{k-r_2+1} \leqslant \frac{r_1}{k-r_1+1} \cdot \frac{k-r_1}{r_1+1}.$$

Thus

$$\rho \leqslant \frac{r_1}{r_1+1} \cdot \frac{k-r_1}{k-r_1+1} < (1-1/k)^2.$$

We have therefore verified that the conditions (9.8) are satisfied.

Next, we have

$$b' = b_{r_2} \leqslant kb_{r_1} \leqslant k^2 b,$$

and since $1 \leqslant r_2 \leqslant k - r_1$, we also have

$$\begin{aligned} b' = b_{r_2} &\geqslant \frac{k - r_2 + 1}{r_2} b_{r_1} \geqslant \frac{r_1 + 1}{k - r_1} b_{r_1} \\ &\geqslant \frac{r_1 + 1}{k - r_1} \cdot \frac{k - r_1 + 1}{r_1} b = \frac{r_1 + 1}{r_1} \cdot \frac{k - r_1 + 1}{k - r_1} b \\ &\geqslant (k/(k-1))^2 b > (1 + 2/k)b. \end{aligned}$$

Thus $(1+2/k)b \leqslant b' \leqslant k^2 b$. Also, since $r' = k - r_2$ and $a' = b_{r_1}$, one finds that

$$\left\lceil \frac{r' + 1}{k - r'} a' \right\rceil = \left\lceil \frac{k - r_2 + 1}{r_2} b_{r_1} \right\rceil = b'.$$

In addition, we have

$$\frac{b'}{b} \geqslant \frac{b_{r_1}/\rho_{r_2}}{b} \geqslant \frac{b/\rho_{r_1}}{\rho_{r_2} b} = \frac{1}{\rho_{r_1}\rho_{r_2}} = \frac{1}{\rho}.$$

Then the conditions (9.9) are satisfied.

Finally, we have

$$a' = b_{r_1} = \left\lceil \frac{k - r_1 + 1}{r_1} b \right\rceil \geqslant \frac{b}{k} \geqslant k\delta\theta$$

and

$$b' \geqslant (1 + 2/k)b \geqslant k^2 \delta\theta.$$

Meanwhile,

$$r'a' = (k - r_2) b_{r_1} < (k - r_2 + 1) b_{r_1},$$

whilst

$$(k - r' + 1)b' = (r_2 + 1) b_{r_2} \geqslant \frac{k - r_2 + 1}{r_2} (r_2 + 1) b_{r_1},$$

so that $r'a' \leqslant (k - r' + 1)b'$. This confirms that the conditions (9.7) are satisfied. All of the conditions (9.7)-(9.9) having been confirmed, the proof of the lemma is complete. □

10. The proof of Theorem 3.1

Our apparatus for the main p-adic concentration argument has now been assembled, in the shape of Lemma 9.3. The lower bound (6.3) may be exploited to show that an initial mean value $\widetilde{K}^1_{\theta,\theta}$ is large. Then, whenever $\widetilde{K}^r_{a,b}$ is large, the estimate (9.6) of Lemma 9.3 shows that a related mean value $\widetilde{K}^{r'}_{a',b'}$ is larger still, and inflated by an additional factor $p^{b\Lambda/(2\rho k)}$. After iteration of this idea, this overabundance of "p-adic energy" blows up to the point that it exceeds even the trivial estimate supplied by Lemma 4.2, delivering a contradiction to the hypothesis that $\Lambda > 0$.

The proof of Theorem 3.1. Throughout, we consider a natural number k and we put $s = k(k+1)/2$. In view of Lemma 5.1, we may suppose that $k \geqslant 2$, and we may also suppose that the conclusion of Theorem 3.1 has already been established for exponents smaller than k.

We aim to show that $\lambda(s,k) \leqslant 0$. We may therefore work throughout under the assumption that $\lambda(s,k) = \Lambda$ with $\Lambda > 0$, and seek a contradiction. Of course, should $\lambda(s,k) \leqslant 0$, then there is nothing to prove. We initiate the iteration with an appeal to Lemma 6.3, which shows that

$$U_{s,k}^B(\mathfrak{a}) \ll p^{s\theta} K_{\theta,\theta}^1(\mathfrak{a}).$$

In view of the assumption (6.3), we deduce from (3.24) that

$$[\![K_{\theta,\theta}^1(\mathfrak{a})]\!]_\Lambda \gg \frac{p^{-s\theta} U_{s,k}^B(\mathfrak{a})}{p^{\Lambda H} U_{s,k}^{B,H}(\mathfrak{a})} \geqslant p^{-s\theta - H\varepsilon}.$$

Our hierarchy (6.2) combines with (6.10), therefore, to ensure that

$$\widetilde{K}_{\theta,\theta}^1 \gg p^{-2s\theta}. \tag{10.1}$$

Next, we apply Lemma 9.3 repeatedly. Put

$$N = \lceil 16sk/\Lambda \rceil. \tag{10.2}$$

Our assumption that B, and hence also H, is sufficiently large in terms of our hierarchy of parameters (6.2) ensures that $2k^{2N+2}\theta < H$. We claim that sequences (a_n), (b_n), (r_n), (ρ_n) may be defined for $0 \leqslant n \leqslant N$ in such a manner that

$$1 \leqslant r_n \leqslant k-1, \quad k^2\delta\theta \leqslant b_n \leqslant k^{2n+2}\theta, \tag{10.3}$$

$$\delta\theta \leqslant a_n \leqslant (k - r_n + 1)b_n/r_n, \tag{10.4}$$

$$0 < \rho_n < (1-1/k)^2, \quad \rho_n b_n \geqslant b_{n-1} \quad (n \geqslant 1), \tag{10.5}$$

and so that

$$\widetilde{K}_{\theta,\theta}^1 \ll (\widetilde{K}_{a_n,b_n}^{r_n})^{\rho_1 \cdots \rho_n} (p^{-\Lambda/(2k)})^{nb_0}. \tag{10.6}$$

We initiate these sequences by putting $a_0 = b_0 = \theta$ and $\rho_0 = r_0 = 1$, so that (10.6) is immediate in the case $n = 0$ from the usual convention that an empty product is 1, whence $\rho_1 \cdots \rho_n = 1$ for $n = 0$.

We now attend to the task of confirming this claim, proceeding by induction on n. Suppose that such has already been confirmed for $0 \leqslant n < m$, with $1 \leqslant m \leqslant N$. Then we have

$$\widetilde{K}_{\theta,\theta}^1 \ll (\widetilde{K}_{a_{m-1},b_{m-1}}^{r_{m-1}})^{\rho_1 \cdots \rho_{m-1}} (p^{-\Lambda/(2k)})^{(m-1)b_0}. \tag{10.7}$$

We estimate $\widetilde{K}_{a_{m-1},b_{m-1}}^{r_{m-1}}$ by appealing to Lemma 9.3. The conditions (10.3)-(10.5) may be assumed to hold with $n = m-1$. Thus, since

$$k^2 b_{m-1} \leqslant k^{2m+2}\theta \leqslant k^{2N+2}\theta < H/2,$$

we see that the hypotheses required to apply Lemma 9.3 are satisfied. Consequently, there exist integers a_m, b_m, r_m, ρ_m having the property that

$$\widetilde{K}_{a_{m-1},b_{m-1}}^{r_{m-1}} \ll (\widetilde{K}_{a_m,b_m}^{r_m})^{\rho_m} p^{-b_{m-1}\Lambda/(2k)}, \tag{10.8}$$

with

$$a_m \geqslant \delta\theta, \quad b_m \geqslant k^2\delta\theta, \quad r_m a_m \leqslant (k - r_m + 1)b_m,$$
$$1 \leqslant r_m \leqslant k-1, \quad 0 < \rho_m < (1-1/k)^2,$$

$$(1+2/k)b_{m-1} \leqslant b_m \leqslant k^2 b_{m-1}, \quad b_m = \left\lceil \frac{r_m+1}{k-r_m} a_m \right\rceil, \quad \rho_m b_m \geqslant b_{m-1}.$$

Since we may suppose from (10.3) in the case $n = m-1$ that $b_{m-1} \leqslant k^{2m}\theta$, we see that $b_m \leqslant k^2 b_{m-1} \leqslant k^{2m+2}\theta$, and so the conditions (10.3)-(10.5) are all met with $n = m$. Moreover, on substituting (10.8) into (10.7), we obtain the bound

$$\widetilde{K}^1_{\theta,\theta} \ll (\widetilde{K}^{r_m}_{a_m,b_m})^{\rho_1 \cdots \rho_m} (p^{-\Lambda/(2k)})^{(m-1)b_0 + \rho_1 \cdots \rho_{m-1} b_{m-1}}.$$

However, the condition (10.5) for $1 \leqslant n \leqslant m-1$ ensures that

$$\rho_1 \cdots \rho_{m-1} b_{m-1} \geqslant \rho_1 \cdots \rho_{m-2} b_{m-2} \geqslant \ldots \geqslant \rho_1 b_1 \geqslant b_0,$$

and so

$$(m-1)b_0 + \rho_1 \cdots \rho_{m-1} b_{m-1} \geqslant m b_0.$$

We therefore conclude that (10.6) holds for $n = m$, and hence our claimed assertion follows by induction for $0 \leqslant n \leqslant N$.

At this point in our argument, we may combine the lower bound (10.1) with the upper bound (10.6) in the case $n = N$. Thus we see that

$$p^{-2s\theta} \ll \widetilde{K}^1_{\theta,\theta} \ll (\widetilde{K}^{r_N}_{a_N,b_N})^{\rho} (p^{-\Lambda/(2k)})^{N\theta}, \tag{10.9}$$

where $\rho = \rho_1 \cdots \rho_N < 1$. On recalling the definition (3.24) and the assumption that $\lambda(s,k) = \Lambda$, the conclusion of Lemma 4.2 shows that

$$[\![K^{r_N}_{a_N,b_N}]\!]_\Lambda \ll p^{H\varepsilon}.$$

Then, again employing the properties of our hierarchy (6.2), we may suppose that $\widetilde{K}^{r_N}_{a_N,b_N} \ll p^\theta$, whence (10.9) delivers the bound

$$p^{-2s\theta} \ll p^{\theta(1-N\Lambda/(2k))},$$

and hence

$$(p^\theta)^{4s} \gg (p^\theta)^{N\Lambda/(2k)}. \tag{10.10}$$

Observe that the definition (6.10) of θ shows that p^θ is sufficiently large in terms of s, k and Λ. Hence, the upper bound (10.10) can hold only when

$$4s \geqslant N\Lambda/(2k).$$

In view of (10.2), we therefore obtain the bound

$$\Lambda \leqslant 8sk/N \leqslant \Lambda/2,$$

which yields a contradiction to the assumption that $\Lambda > 0$.

We are therefore forced to conclude that Λ cannot be positive, whence $\lambda(s,k) = 0$. This completes the proof of Theorem 3.1 for the exponent k, and the theorem follows in full by induction on k. □

Corollary 3.2 follows from Theorem 3.1 by simply interpreting the definitions (3.12) and (3.13) of $\lambda^*(s,\theta;\tau)$ and $\lambda(s,\theta)$. Thus, since $\lambda(s,k) = 0$, we may suppose that for every $\varepsilon > 0$, whenever $\tau > 0$ is sufficiently small, one has $\lambda^*(s,k;\tau) < \varepsilon$. But for each $\varepsilon > 0$, one has

$$U^B_{s,k}(\mathfrak{a}) \ll p^{(\lambda^*(s,k;\tau)+\varepsilon)H} U^{B,H}_{s,k}(\mathfrak{a})$$

for all sequences $(\mathfrak{a}_n) \in \mathbb{D}$, all $\boldsymbol{\varphi} \in \Phi_\tau(B)$ and all large enough values of B. The corresponding conclusion for the zero sequence $(\mathfrak{a}_n) \in \mathbb{D}_0 \setminus \mathbb{D}$ is, of course, trivial.

The inquisitive reader might wonder at what point in the proof of Theorem 3.1 and Corollary 3.2 did we find ourselves limited to the situation in which $H = \lceil B/k \rceil$, rather than allowing the possibility that H might exceed $\lceil B/k \rceil$. This is a little subtle, since at face value our argument does not involve estimates for $K^{r,\boldsymbol{\varphi},\nu}_{a,b,c}(\mathfrak{a})$ with a and b close in size to B/k. However, in the proof of Lemma 4.1, in equation (4.9), one encounters a situation wherein certain congruences would be trivially satisfied were h to exceed B/k. This failure of independence amongst the congruences would compromise our estimates, and implicitly generate associated difficulties in §8. Since we are imposing the condition $h \leqslant (1-\delta)H$ in Lemma 4.1, one finds it necessary to restrict H to be no larger than about B/k.

11. Solutions of congruences in short intervals

Rather than embark at once on the proof of Theorem 1.1 and its corollaries, we spend some time in this section on a more immediate application of Theorem 3.1 to the topic of solutions of congruences in short intervals. We are interested in rational functions $\chi_j \in \mathbb{Q}(t)$ $(1 \leqslant j \leqslant k)$, and the number of integral solutions of systems of congruences of the shape

$$\sum_{i=1}^{s} \chi_j(x_i) \equiv \sum_{i=1}^{s} \chi_j(y_i) \pmod{p^B} \quad (1 \leqslant j \leqslant k), \tag{11.1}$$

with $X < x_i, y_i \leqslant X + Y$. Writing $\chi_j(t) = \varphi_j(t)/\gamma_j(t)$ for suitable polynomials $\varphi_j, \gamma_j \in \mathbb{Z}[t]$ with $(\varphi_j(t), \gamma_j(t)) = 1$, it is apparent that it is sensible to exclude choices for the variables x with $\gamma_j(x) \equiv 0 \pmod{p}$. With such choices for x excluded, there is a multiplicative inverse for $\gamma_j(x)$, say $\gamma_j(x)^{-1}$ modulo p^B, and this counting problem makes sense. We are also able to make sense of the Wronskian $W(t;\boldsymbol{\chi})$ defined by (1.1) in these circumstances. Indeed, as a rational number A/Q in lowest terms, the Wronskian $W(x;\boldsymbol{\chi})$ has denominator Q not divisible by p in the situation under discussion. If Q^{-1} denotes an integer defining the multiplicative inverse of Q modulo p^B, then we shall always replace $W(x;\boldsymbol{\chi})$ by the integer AQ^{-1}.

We obtain strong estimates for the number of solutions of the system (11.1) when $s \leqslant k(k+1)/2$ and $Y \leqslant p^{B/k}$, and we restrict to solutions with both denominators γ_j and the Wronskian non-vanishing modulo p.

Theorem 11.1. *When $k \in \mathbb{N}$ and $1 \leqslant j \leqslant k$, suppose that $\chi_j \in \mathbb{Q}(t)$ is a rational function with $\chi_j = \varphi_j/\gamma_j$ for suitable polynomials $\varphi_j, \gamma_j \in \mathbb{Z}[t]$ satisfying $(\varphi_j, \gamma_j) = 1$. Let s be a positive number with $s \leqslant k(k+1)/2$, and let p be a prime number with $p > k$. Also, suppose that $(\mathfrak{a}_n)_{n\in\mathbb{Z}}$ is a sequence of complex numbers. Define*

$$h(\boldsymbol{\alpha}; X, Y) = \sum_{X<n\leqslant X+Y}^{*} \mathfrak{a}_n e(\alpha_1\chi_1(n) + \ldots + \alpha_k\chi_k(n)),$$

where the summation is restricted by the conditions

$$(W(n;\boldsymbol{\chi}),p)=1 \quad \textit{and} \quad (\gamma_j(n),p)=1 \quad (1\leqslant j\leqslant k). \tag{11.2}$$

Then whenever $\varepsilon>0$ and B is sufficiently large in terms of ε and k, one has

$$\oint_{p^B}|h(\boldsymbol{\alpha};X,Y)|^{2s}\,\mathrm{d}\boldsymbol{\alpha}\ll p^{B\varepsilon}\bigg(1+\frac{Y}{p^{B/k}}\bigg)^s\bigg(\sum_{X<n\leqslant X+Y}|\mathfrak{a}_n|^2\bigg)^s. \tag{11.3}$$

Corollary 11.2. *With the hypotheses of Theorem 11.1, denote by $N_B(X,Y)$ the number of integral solutions of the system of congruences (11.1) subject for $1\leqslant i\leqslant s$ to the conditions $X<x_i,y_i\leqslant X+Y$ and*

$$(W(x_i;\boldsymbol{\chi})W(y_i;\boldsymbol{\chi}),p)=1 \quad \textit{and} \quad (\gamma_j(x_i)\gamma_j(y_i),p)=1 \quad (1\leqslant j\leqslant k).$$

Then whenever $\varepsilon>0$ and B is sufficiently large in terms of ε and k, one has

$$N_B(X,Y)\ll p^{B\varepsilon}(Y+1)^s\bigg(\frac{Y}{p^{B/k}}+1\bigg)^s.$$

In particular, when $p^B\geqslant Y^k$, one has $N_B(X,Y)\ll p^{B\varepsilon}(Y+1)^s$.

It is apparent that, in general, the diagonal solutions in (11.1) make a contribution of order Y^s to $N_B(X,Y)$, so the conclusion of Corollary 11.2 is essentially best possible.

The proof of Theorem 11.1. Our goal is to transform the mean value on the left hand side of (11.3) into one amenable to Theorem 3.1. We may plainly suppose that the sequence $(\mathfrak{a}_n)$ satisfies the condition that $\mathfrak{a}_n=0$ for $n\leqslant X$ and for $n>X+Y$. We are also at liberty, moreover, to assume that $\mathfrak{a}_n=0$ unless the conditions (11.2) all hold. Recalling the notation of §3, we may then define

$$F_p(\boldsymbol{\alpha})=\rho_0(\mathfrak{a})^{-1}\sum_{n\in\mathbb{Z}}\mathfrak{a}_n e(\psi(n;\boldsymbol{\alpha})),$$

where we now write

$$\psi(n;\boldsymbol{\alpha})=\alpha_1\chi_1(n)+\ldots+\alpha_k\chi_k(n).$$

With this notation, the claimed bound (11.3) translates into the assertion that

$$\oint_{p^B}|F_p(\boldsymbol{\alpha})|^{2s}\,\mathrm{d}\boldsymbol{\alpha}\ll p^{B\varepsilon}\bigg(\frac{Y}{p^{B/k}}+1\bigg)^s, \tag{11.4}$$

and it is this that we now seek to establish. An application of Hölder's inequality delivers the conclusion (11.4) for $0<s\leqslant k(k+1)/2$ from that in the special case $s=k(k+1)/2$. We therefore restrict attention henceforth to the case $s=k(k+1)/2$.

Some preparation is required prior to the proof of the estimate (11.4) via Corollary 3.2. Let $\tau>0$ be sufficiently small in terms of s and k. Then we may suppose that B is sufficiently large in terms of τ, as well as s, k and ε. We put $c=\lceil\tau B\rceil$. Next we sort the implicit summation in $F_p(\boldsymbol{\alpha})$ into arithmetic progressions modulo p^c. In view of our assumptions on the sequence

$(\mathfrak{a}_n)$ implied by the conditions (11.2), we may suppose that $\mathfrak{a}_n = 0$ whenever $n \equiv \xi \pmod{p^c}$ and

$$W(\xi;\boldsymbol{\chi})\gamma_1(\xi)\cdots\gamma_k(\xi) \equiv 0 \pmod{p}.$$

With this assumption, one finds that

$$F_p(\boldsymbol{\alpha}) = \rho_0(\mathfrak{a})^{-1} \sum_{\xi \bmod p^c} \rho_c(\xi)\mathfrak{f}_c(\boldsymbol{\alpha};\xi),$$

and hence Lemma 6.2 delivers the bound

$$\rho_0(\mathfrak{a})^2|F_p(\boldsymbol{\alpha})|^{2s} \leqslant p^{sc} \sum_{\xi \bmod p^c} \rho_c(\xi)^2|\mathfrak{f}_c(\boldsymbol{\alpha};\xi)|^{2s}.$$

Thus we obtain the estimate

$$\oint_{p^B} |F_p(\boldsymbol{\alpha})|^{2s}\,\mathrm{d}\boldsymbol{\alpha} \leqslant p^{sc}\rho_0(\mathfrak{a})^{-2} \sum_{\xi \bmod p^c} \rho_c(\xi)^2 I_p(\xi), \tag{11.5}$$

where

$$I_p(\xi) = \oint_{p^B} |\mathfrak{f}_c(\boldsymbol{\alpha};\xi)|^{2s}\,\mathrm{d}\boldsymbol{\alpha}.$$

The mean value $I_p(\xi)$ counts the integral solutions $\mathbf{y}, \mathbf{z}$ of the system of congruences

$$\sum_{i=1}^{s} (\chi_j(p^c y_i + \xi) - \chi_j(p^c z_i + \xi)) \equiv 0 \pmod{p^B} \quad (1 \leqslant j \leqslant k), \tag{11.6}$$

with each solution being counted with weight

$$\rho_c(\xi)^{-2s} \prod_{i=1}^{s} \mathfrak{a}_{p^c y_i + \xi}\overline{\mathfrak{a}}_{p^c z_i + \xi}. \tag{11.7}$$

Here, we note that we may restrict attention to the situation in which

$$W(\xi;\boldsymbol{\chi})\gamma_1(\xi)\cdots\gamma_k(\xi) \not\equiv 0 \pmod{p} \quad \text{and} \quad \rho_c(\xi) > 0.$$

In particular, it follows from (3.8) and (11.5) that

$$\oint_{p^B} |F_p(\boldsymbol{\alpha})|^{2s}\,\mathrm{d}\boldsymbol{\alpha} \leqslant p^{sc} U_{s,k}^{B,c}(\mathfrak{a}).$$

We reinterpret the system (11.6) by applying Taylor's theorem to expand the rational functions $\chi_j(p^c t + \xi)$. When $1 \leqslant j \leqslant k$, we find that for suitable polynomials $\Phi_j \in \mathbb{Z}[t]$, one has

$$\chi_j(p^c t + \xi) - \chi_j(\xi) \equiv \sum_{l=1}^{k} \omega_{lj}(p^c t)^l + (p^c t)^{k+1}\Phi_j(p^c t) \pmod{p^B},$$

in which the integral coefficients ω_{lj} are defined by taking

$$\omega_{lj} = (l!)^{-1}\chi_j^{(l)}(\xi) \quad (1 \leqslant l, j \leqslant k).$$

A few words of explanation are in order here. First, since $\varphi_j = \gamma_j \chi_j$, it follows from the differentiation rule of Leibniz that

$$\varphi_j^{(l)} = \sum_{m=0}^{l} \binom{l}{m} \gamma_j^{(m)} \chi_j^{(l-m)},$$

whence one obtains the iterative relation

$$\gamma_j(\xi) \frac{\chi_j^{(l)}(\xi)}{l!} = \frac{\varphi_j^{(l)}(\xi)}{l!} - \sum_{m=1}^{l} \binom{l}{m} \left(\frac{\gamma_j^{(m)}(\xi)}{m!} \right) \left(\frac{\chi_j^{(l-m)}(\xi)}{(l-m)!} \right).$$

But since $\gamma_j, \varphi_j \in \mathbb{Z}[t]$, one finds that $m!$ divides every coefficient of $\gamma_j^{(m)}(t)$, and likewise $l!$ divides every coefficient of $\varphi_j^{(l)}(t)$. In addition, we may suppose that $\gamma_j(\xi) \not\equiv 0 \pmod{p}$. An inductive argument therefore conveys us from this iterative relation to the conclusion that $\mathrm{ord}_p(l!) \leqslant \mathrm{ord}_p(\chi_j^{(l)}(\xi))$ for all non-negative integers l. By multiplying through by appropriate multiplicative inverses modulo p^B, we may thus suppose that $(l!)^{-1}\chi_j^{(l)}(\xi)$ is an integer for $l \geqslant 0$. Second, the expansion of $\chi_j(p^c t + \xi)$ might be expected to be a non-terminating infinite series. However, the terms $(l!)^{-1}\chi_j^{(l)}(\xi)(p^c t)^l$ are necessarily congruent to 0 modulo p^B whenever l is sufficiently large in terms of B.

The determinant of the coefficient matrix $\Omega = (\omega_{lj})_{1 \leqslant l,j \leqslant k}$ is given by the formula

$$\det(\Omega) = W(\xi; \boldsymbol{\chi}) \biggl(\prod_{l=1}^{k} l! \biggr)^{-1}.$$

Since we may suppose that $p > k$, the hypothesis $(W(\xi; \boldsymbol{\chi}), p) = 1$ permits us to conclude that $(\det(\Omega), p) = 1$, whence Ω possesses a multiplicative inverse Ω^{-1} modulo p^B having integral coefficients. We now replace $\boldsymbol{\chi}$ by $\Omega^{-1}\boldsymbol{\chi}$ and $\boldsymbol{\Phi}$ by $\Omega^{-1}\boldsymbol{\Phi}$. This amounts to taking suitable integral linear combinations of the congruences comprising (11.6). In this way, we see that there is no loss of generality in supposing that the coefficient matrix Ω is equal to I_k. Hence, there exist polynomials $\Xi_j \in \mathbb{Z}[t]$ having the property that whenever the system (11.6) is satisfied, then

$$\sum_{i=1}^{s} (p^c)^j \left(\Psi_j(y_i) - \Psi_j(z_i) \right) \equiv 0 \pmod{p^B} \quad (1 \leqslant j \leqslant k),$$

in which $\Psi_j(t) = t^j + p^c \Xi_j(t)$. In particular, the system of polynomials $\boldsymbol{\Psi}$ is p^c-spaced.

The discussion of the previous paragraph shows that in any solution $\mathbf{y}, \mathbf{z}$ of the system (11.6), counted with weight (11.7), one has the additional constraints

$$\sum_{i=1}^{s} \left(\Psi_j(y_i) - \Psi_j(z_i) \right) \equiv 0 \pmod{p^{B-kc}} \quad (1 \leqslant j \leqslant k).$$

Define the coefficients

$$\mathfrak{c}_y(\boldsymbol{\alpha}) = \mathfrak{a}_{p^c y + \xi} e(\psi(p^c y + \xi; \boldsymbol{\alpha})).$$

Also, write

$$\mathfrak{g}_{\mathfrak{c}}(\boldsymbol{\alpha},\boldsymbol{\beta})=\rho_0(\mathfrak{c})^{-1}\sum_{y\in\mathbb{Z}}\mathfrak{c}_y(\boldsymbol{\alpha})e(\beta_1\Psi_1(y)+\ldots+\beta_k\Psi_k(y)),$$

and define the mean value

$$J(\boldsymbol{\alpha})=\oint_{p^{B-kc}}|\mathfrak{g}_{\mathfrak{c}}(\boldsymbol{\alpha},\boldsymbol{\beta})|^{2s}\,\mathrm{d}\boldsymbol{\beta}.$$

Note that $\rho_0(\mathfrak{c})=\rho_c(\xi;\mathfrak{a})$. Then, just as in the argument leading to (7.18) above, one sees that

$$I_p(\xi)=\oint_{p^B}|\mathfrak{g}_{\mathfrak{c}}(\boldsymbol{\alpha};\mathbf{0})|^{2s}\,\mathrm{d}\boldsymbol{\alpha}=\oint_{p^B}J(\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha}. \tag{11.8}$$

Observe that, when written using the notation defined in (3.6), one has $J(\boldsymbol{\alpha})=U_{s,k}^{B-kc,\boldsymbol{\Psi}}(\mathfrak{c})$. Moreover, the system $\boldsymbol{\Psi}$ is p^c-spaced and $B-kc$ is sufficiently large in terms of k, τ and ε. Then Corollary 3.2 yields the bound

$$J(\boldsymbol{\alpha})\ll p^{B\varepsilon}U_{s,k}^{B-kc,H,\boldsymbol{\Psi}}(\mathfrak{c}),$$

where $H=\lceil B/k\rceil-c$. On substituting this bound into (11.8), and thence into (11.5), we infer that

$$\oint_{p^B}|F_p(\boldsymbol{\alpha})|^{2s}\,\mathrm{d}\boldsymbol{\alpha}\ll p^{sc+B\varepsilon}\rho_0(\mathfrak{a})^{-2}\sum_{\xi\bmod p^c}\rho_c(\xi)^2U_{s,k}^{B-kc,H,\boldsymbol{\Psi}}(\mathfrak{c}). \tag{11.9}$$

Temporarily, we abbreviate $\mathfrak{f}_H(\boldsymbol{\beta};\eta;\mathfrak{c};\boldsymbol{\Psi})$ to $\mathfrak{f}_H(\boldsymbol{\beta};\eta)$. Thus, we have

$$\mathfrak{f}_H(\boldsymbol{\beta};\eta)=\rho_H(\eta;\mathfrak{c})^{-1}\sum_{y\equiv\eta\ (\mathrm{mod}\ p^H)}\mathfrak{c}_y(\boldsymbol{\alpha})e(\beta_1\Psi_1(y)+\ldots+\beta_k\Psi_k(y)),$$

wherein the coefficients $\mathfrak{c}_y(\boldsymbol{\alpha})$ are 0 whenever $p^cy+\xi\leqslant X$ or $p^cy+\xi>X+Y$. It follows via Cauchy's inequality, therefore, that

$$\rho_H(\eta;\mathfrak{c})^2|\mathfrak{f}_H(\boldsymbol{\beta};\eta)|^2\leqslant\Bigl(\sum_{\substack{y\equiv\eta\ (\mathrm{mod}\ p^H)\\ X<p^cy+\xi\leqslant X+Y}}1\Bigr)\sum_{y\equiv\eta\ (\mathrm{mod}\ p^H)}|\mathfrak{c}_y(\boldsymbol{\alpha})|^2.$$

Moreover, one has

$$\rho_H(\eta;\mathfrak{c})^2=\sum_{y\equiv\eta\ (\mathrm{mod}\ p^H)}|\mathfrak{c}_y(\boldsymbol{\alpha})|^2,$$

so that

$$|\mathfrak{f}_H(\boldsymbol{\beta};\eta)|^2\leqslant 1+Y/p^{c+H}.$$

We therefore infer from (3.8) that

$$U_{s,k}^{B-kc,H,\boldsymbol{\Psi}}(\mathfrak{c})\ll(1+Y/p^{c+H})^s.$$

Since $c+H=\lceil B/k\rceil$, we conclude from (11.9) that

$$\begin{aligned}\oint_{p^B}|F_p(\boldsymbol{\alpha})|^{2s}\,\mathrm{d}\boldsymbol{\alpha}&\ll p^{sc+B\varepsilon}(1+Y/p^{B/k})^s\rho_0(\mathfrak{a})^{-2}\sum_{\xi\bmod p^c}\rho_c(\xi)^2\\&\ll p^{(2s\tau+\varepsilon)B}(1+Y/p^{B/k})^s.\end{aligned}$$

We recall that τ was chosen sufficiently small in terms of s and k. Thus, for each positive number δ, we have

$$\oint_{p^B} |F_p(\boldsymbol{\alpha})|^{2s}\,\mathrm{d}\boldsymbol{\alpha} \ll p^{B\delta}(1+Y/p^{B/k})^s,$$

so that the estimate (11.4) does indeed hold. This completes the proof of the theorem. □

Corollary 11.2 is immediate from Theorem 11.1 via orthogonality.

12. Mean values of exponential sums

We now explain how Theorem 3.1 and Corollary 3.2 may be applied to bound mean values of exponential sums. In particular, we establish Theorem 1.1 and its corollaries. Much of the necessary work is already accomplished in the shape of Theorem 11.1.

The proof of Theorem 1.1. Let $\varphi_j \in \mathbb{Z}[t]$ $(1 \leqslant j \leqslant k)$ be polynomials with non-vanishing Wronskian $W(t;\boldsymbol{\varphi})$, and let $\varepsilon > 0$ be a small positive number. Let $\mathcal{Z}$ denote the set of integral zeros of $W(t;\boldsymbol{\varphi})$, and let X be sufficiently large in terms of $\boldsymbol{\varphi}$, s, k and ε. We note that

$$\mathrm{card}(\mathcal{Z}) \leqslant \deg(W(t;\boldsymbol{\varphi})) \ll 1.$$

We suppose that $s = k(k+1)/2$. Finally, when $(\mathfrak{a}_n)_{n\in\mathbb{Z}}$ is a sequence of complex numbers, we write

$$F(\boldsymbol{\alpha};X) = \rho_0^{-1} \sum_{|n|\leqslant X} \mathfrak{a}_n e(\psi(n;\boldsymbol{\alpha}))$$

and

$$F_0(\boldsymbol{\alpha};X) = \rho_0^{-1} \sum_{\substack{|n|\leqslant X\\ n\notin\mathcal{Z}}} \mathfrak{a}_n e(\psi(n;\boldsymbol{\alpha})),$$

where $\psi(n;\boldsymbol{\alpha})$ is defined as in (3.3). In the argument to come, there will be no loss of generality in supposing that $\mathfrak{a}_n = 0$ for $|n| > X$.

As a consequence of Cauchy's inequality, one has

$$\Big| \sum_{\substack{|n|\leqslant X\\ n\in\mathcal{Z}}} \mathfrak{a}_n e(\psi(n;\boldsymbol{\alpha})) \Big|^2 \leqslant \mathrm{card}(\mathcal{Z}) \sum_{|n|\leqslant X} |\mathfrak{a}_n|^2 \leqslant \rho_0^2\,\mathrm{card}(\mathcal{Z}),$$

whence

$$|F(\boldsymbol{\alpha};X)| \ll 1 + |F_0(\boldsymbol{\alpha};X)|.$$

Thus, one has

$$|F(\boldsymbol{\alpha};X)|^{2s} \ll 1 + |F_0(\boldsymbol{\alpha};X)|^{2s},$$

so that

$$\oint |F(\boldsymbol{\alpha};X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha} \ll 1 + \oint |F_0(\boldsymbol{\alpha};X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha}. \tag{12.1}$$

In this way, we discern that it suffices to restrict attention to situations in which the underlying variables possess non-vanishing Wronskians. By orthogonality,

the mean value on the right hand side of (12.1) counts the number of integral solutions of the system of equations (1.4) with $|\mathbf{x}|, |\mathbf{y}| \leqslant X$ and $x_i, y_i \notin \mathcal{Z}$, and with each solution $\mathbf{x}, \mathbf{y}$ being counted with weight

$$\rho_0^{-2s} \prod_{i=1}^{s} \mathfrak{a}_{x_i} \overline{\mathfrak{a}}_{y_i}. \tag{12.2}$$

We impose a non-vanishing condition modulo p on the Wronskian in each of these solutions, for a suitable prime number p. Given a solution $\mathbf{x}, \mathbf{y}$ of (1.4) of the type in question, the integer

$$\Xi(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^{s} W(x_i; \boldsymbol{\varphi}) W(y_i; \boldsymbol{\varphi})$$

is non-zero. Moreover, one has $|\Xi(\mathbf{x}, \mathbf{y})| \leqslant CX^D$, for some $C > 0$ depending at most on s, k and the coefficients of $\boldsymbol{\varphi}$, and D a positive integer with

$$D \leqslant 2s \sum_{j=1}^{k} \deg(\varphi_j).$$

Let $\mathcal{P}$ denote the set of prime numbers p with

$$(\log X)^2 < p \leqslant 3(\log X)^2.$$

Thus, when X is large, it is a consequence of the prime number theorem that

$$\prod_{p \in \mathcal{P}} p > (\log X)^{(\log X)^2/\log\log X} > CX^D.$$

We therefore deduce that for each solution $\mathbf{x}, \mathbf{y}$ of (1.4) counted by the integral on the right hand side of (12.1), there exists $p \in \mathcal{P}$ with

$$\prod_{i=1}^{s} W(x_i; \boldsymbol{\varphi}) W(y_i; \boldsymbol{\varphi}) \not\equiv 0 \pmod{p}.$$

In particular, one has

$$\oint |F_0(\boldsymbol{\alpha}; X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \leqslant \sum_{p \in \mathcal{P}} \oint |F_p(\boldsymbol{\alpha}; X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha},$$

where

$$F_p(\boldsymbol{\alpha}; X) = \rho_0^{-1} \sum_{W(n;\boldsymbol{\varphi}) \not\equiv 0 \pmod{p}} |\mathfrak{a}_n| e(\psi(n; \boldsymbol{\alpha})). \tag{12.3}$$

Thus, we conclude from (12.1) that

$$\oint |F(\boldsymbol{\alpha}; X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \ll 1 + (\log X)^2 \max_{p \in \mathcal{P}} \oint |F_p(\boldsymbol{\alpha}; X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha}. \tag{12.4}$$

Our goal is to establish that for each fixed prime $p \in \mathcal{P}$, one has

$$\oint |F_p(\boldsymbol{\alpha}; X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \ll X^{\varepsilon/2}. \tag{12.5}$$

It follows by substituting this estimate into (12.4) that

$$\oint |F(\boldsymbol{\alpha}; X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \ll 1 + X^{\varepsilon/2} (\log X)^2 \ll X^{\varepsilon},$$

whence

$$\oint \bigg| \sum_{|n| \leqslant X} \mathfrak{a}_n e(\psi(n; \boldsymbol{\alpha})) \bigg|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \ll X^{\varepsilon} \rho_0^{2s} = X^{\varepsilon} \bigg(\sum_{|n| \leqslant X} |\mathfrak{a}_n|^2 \bigg)^s. \tag{12.6}$$

This establishes the first conclusion (1.2) of Theorem 1.1 in the special case $s = k(k+1)/2$. Of course, the desired conclusion for smaller values of s follows by applying Hölder's inequality. Meanwhile, the final conclusion (1.3) of Theorem 1.1 follows from (1.2) by merely specialising the sequence $(\mathfrak{a}_n)$ to be (1).

We focus now on the proof of the estimate (12.5), and this involves preparation for the application of Theorem 11.1. In our application of Theorem 11.1 we take $\gamma_j = 1$, so that $\chi_j = \varphi_j$ $(1 \leqslant j \leqslant k)$. In view of the definition (12.3), we may suppose that $\mathfrak{a}_n$ is real with $\mathfrak{a}_n \geqslant 0$ for each n. Also, we take

$$B = \bigg\lceil \frac{k \log X}{\log p} \bigg\rceil.$$

Thus $X \leqslant p^{B/k} \leqslant pX$.

The mean value on the left hand side of (12.5) counts the integral solutions of the system of equations (1.4) with $|\mathbf{x}|, |\mathbf{y}| \leqslant X$ satisfying

$$W(x_i; \boldsymbol{\varphi}) W(y_i; \boldsymbol{\varphi}) \not\equiv 0 \pmod{p} \quad (1 \leqslant i \leqslant s),$$

and in which each solution is counted with weight (12.2). Since these weights are now assumed to be non-negative, one sees that an upper bound for this weighted number of solutions is given by

$$\oint_{p^B} |F_p(\boldsymbol{\alpha}; X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha},$$

for by orthogonality this mean value counts the integral solutions $\mathbf{x}, \mathbf{y}$ of the system of congruences

$$\sum_{i=1}^{s} (\varphi_j(x_i) - \varphi_j(y_i)) \equiv 0 \pmod{p^B} \quad (1 \leqslant j \leqslant k),$$

with $\mathbf{x}, \mathbf{y}$ satisfying the same attendant conditions, and again counted with weight (12.2).

By Theorem 11.1, one has

$$\oint_{p^B} |F_p(\boldsymbol{\alpha}; X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \ll \rho_0^{-2s} p^{B\varepsilon/(4k)} \bigg(\frac{X}{p^{B/k}} + 1 \bigg)^s \bigg(\sum_{|n| \leqslant X} |\mathfrak{a}_n|^2 \bigg)^s \ll p^{B\varepsilon/(4k)}.$$

Since $p^{B\varepsilon} \ll X^{(k+1)\varepsilon}$, we confirm the estimate (12.5). Thus we deduce from (12.4) that one has the bound (12.6). The proof of Theorem 1.1, as previously discussed, is now complete. □

The proof of Corollary 1.2 involves only a computation involving the Wronskian.

The proof of Corollary 1.2. In order to establish Corollary 1.2, we have merely to note that when $1 \leqslant d_1 < d_2 < \ldots < d_k$ and $\varphi_j(t) = t^{d_j}$ $(1 \leqslant j \leqslant k)$, then the Wronskian $W(t;\boldsymbol{\varphi})$ is non-zero as a polynomial. In order to see this, observe that

$$W(t;\boldsymbol{\varphi}) = \det\big(d_j(d_j-1)\cdots(d_j-i+1)t^{d_j-i}\big)_{1\leqslant i,j\leqslant k}.$$

Every monomial in this determinant is an integer multiple of t^D, where

$$D = \bigg(\sum_{j=1}^k d_j\bigg) - k(k+1)/2 = \sum_{j=1}^k (d_j - j).$$

Thus $W(t;\boldsymbol{\varphi}) = t^D \det(\Omega)$, where

$$\Omega = (d_j(d_j-1)\cdots(d_j-i+1))_{1\leqslant i,j\leqslant k}.$$

By taking appropriate linear combinations of the rows of the matrix Ω, one sees that

$$\det(\Omega) = \det(d_j^i)_{1\leqslant i,j\leqslant t} = d_1\cdots d_k \prod_{1\leqslant i<j\leqslant k}(d_i - d_j).$$

Thus we see that $\det(\Omega) \neq 0$, and hence $W(t;\boldsymbol{\varphi}) \neq 0$. We therefore conclude from Theorem 1.1 that when $0 < s \leqslant k(k+1)/2$, one has

$$\oint \bigg| \sum_{1\leqslant x\leqslant X} e(\alpha_1 x^{d_1} + \ldots + \alpha_k x^{d_k})\bigg|^{2s} \mathrm{d}\boldsymbol{\alpha} \ll X^{s+\varepsilon}.$$

This completes the proof of Corollary 1.2. □

We address the proof of Corollary 1.4 before moving on to consider Corollary 1.3, this being logically speaking a more direct course of action.

The proof of Corollary 1.4. Write

$$\mathfrak{h}(\boldsymbol{\alpha}) = \sum_{|n|\leqslant X} \mathfrak{a}_n e(n\alpha_1 + \ldots + n^k\alpha_k).$$

When $\varphi_j(t) = t^j$ $(1 \leqslant j \leqslant k)$, one finds that the Wronskian determinant is triangular, whence

$$W(t;\varphi) = \prod_{j=1}^k j! \neq 0.$$

Thus the conditions required to apply Theorem 1.1 hold, and one obtains the bound

$$\oint |\mathfrak{h}(\boldsymbol{\alpha})|^{2s}\,\mathrm{d}\boldsymbol{\alpha} \ll X^{\varepsilon}\bigg(\sum_{|n|\leqslant X}|\mathfrak{a}_n|^2\bigg)^s \tag{12.7}$$

for $0 < s \leqslant k(k+1)/2$. Meanwhile, an application of Cauchy's inequality shows that

$$|\mathfrak{h}(\boldsymbol{\alpha})|^2 \leqslant (2X+1)\sum_{|n|\leqslant X}|\mathfrak{a}_n|^2.$$

Thus, when X is large, we deduce from the special case $s = k(k+1)/2$ of (12.7) that when $t > k(k+1)/2$, then

$$\oint |\mathfrak{h}(\boldsymbol{\alpha})|^{2t}\,\mathrm{d}\boldsymbol{\alpha} \ll X^{t-k(k+1)/2}\Big(\sum_{|n|\leqslant X}|\mathfrak{a}_n|^2\Big)^{t-k(k+1)/2}\oint |\mathfrak{h}(\boldsymbol{\alpha})|^{k(k+1)}\,\mathrm{d}\boldsymbol{\alpha}$$
$$\ll X^{t-k(k+1)/2+\varepsilon}\Big(\sum_{|n|\leqslant X}|\mathfrak{a}_n|^2\Big)^{t},$$

and the proof of the corollary is complete. □

The proof of Corollary 1.3. For each $\varepsilon > 0$, the estimate

$$J_{s,k}(X) \ll X^{\varepsilon}(X^s + X^{2s-k(k+1)/2})$$

is immediate from the special case of Corollary 1.4 in which $(\mathfrak{a}_n) = (1)$. Meanwhile, the asymptotic formula claimed in the corollary for $s > k(k+1)/2$ follows from this estimate by the standard literature in Vinogradov's mean value theorem, and permits the factor X^{ε} to be omitted from this upper bound when $s > k(k+1)/2$. The asymptotic formula for $J_{s,k}(X)$ asserted in Corollary 1.3 is a well-known consequence of the main conjecture in Vinogradov's mean value theorem. We briefly outline how to apply the circle method to prove this formula.

Write $L = X^{1/(4k)}$. Then, when $1 \leqslant q \leqslant L$, $1 \leqslant a_j \leqslant q$ $(1 \leqslant j \leqslant k)$ and $(q, a_1, \dots, a_k) = 1$, define the major arc $\mathfrak{M}(q, \mathbf{a})$ by

$$\mathfrak{M}(q, \mathbf{a}) = \{\boldsymbol{\alpha} \in [0,1)^k : |\alpha_j - a_j/q| \leqslant LX^{-j}\ (1 \leqslant j \leqslant k)\}.$$

The arcs $\mathfrak{M}(q, \mathbf{a})$ are disjoint, as is easily verified. Let $\mathfrak{M}$ denote their union, and put $\mathfrak{m} = [0,1)^k \setminus \mathfrak{M}$.

Write

$$f(\boldsymbol{\alpha}; X) = \sum_{1\leqslant x\leqslant X} e(\alpha_1 x + \ldots + \alpha_k x^k).$$

Also, when $\boldsymbol{\alpha} \in \mathfrak{M}(q, \mathbf{a}) \subseteq \mathfrak{M}$, put

$$V(\boldsymbol{\alpha}; q, \mathbf{a}) = q^{-1}S(q, \mathbf{a})I(\boldsymbol{\alpha} - \mathbf{a}/q; X),$$

where

$$S(q, \mathbf{a}) = \sum_{r=1}^{q} e((a_1 r + \ldots + a_k r^k)/q)$$

and

$$I(\boldsymbol{\beta}; X) = \int_0^X e(\beta_1\gamma + \ldots + \beta_k\gamma^k)\,\mathrm{d}\gamma.$$

We then define the function $V(\boldsymbol{\alpha})$ to be $V(\boldsymbol{\alpha}; q, a)$ when $\boldsymbol{\alpha} \in \mathfrak{M}(q, \mathbf{a}) \subseteq \mathfrak{M}$, and to be 0 otherwise.

The contribution of the minor arcs $\mathfrak{m}$ is easily estimated. Thus, as in [59, equation (7.1)] (based on [47, §9]), we find that for each $\varepsilon > 0$ one has

$$\sup_{\boldsymbol{\alpha}\in\mathfrak{m}} |f(\boldsymbol{\alpha}; X)| \ll X^{1-\tau+\varepsilon},$$

where $\tau^{-1} = 8k^2$. In this way, when $2s > k(k+1)$, one finds that

$$\int_{\mathfrak{m}} |f(\boldsymbol{\alpha}; X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \ll \Big(\sup_{\boldsymbol{\alpha} \in \mathfrak{m}} |f(\boldsymbol{\alpha}; X)| \Big)^{2s-k(k+1)} \oint |f(\boldsymbol{\alpha}; X)|^{k(k+1)} \, \mathrm{d}\boldsymbol{\alpha} \\ \ll X^{2s-k(k+1)/2+2\varepsilon-(2s-k(k+1))\tau}.$$

Thus

$$\int_{\mathfrak{m}} |f(\boldsymbol{\alpha}; X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} = o(X^{2s-k(k+1)/2}), \tag{12.8}$$

provided that we take ε sufficiently small in terms of s and k.

On the other hand, as in the proof of [59, Lemma 7.1], one finds that when $\boldsymbol{\alpha} \in \mathfrak{M}(q, \mathbf{a}) \subseteq \mathfrak{M}$, one has

$$f(\boldsymbol{\alpha}; X) - V(\boldsymbol{\alpha}; q, \mathbf{a}) \ll L^2.$$

Using the decomposition

$$z\overline{z} - w\overline{w} = (z-w)\overline{z} + w(\overline{z} - \overline{w}),$$

it follows that

$$|f(\boldsymbol{\alpha}; X)|^2 - |V(\boldsymbol{\alpha}; q, \mathbf{a})|^2 \ll L^2 X.$$

Hence, as a consequence of the mean value theorem, one obtains the bound

$$|f(\boldsymbol{\alpha}; X)|^{2s} - |V(\boldsymbol{\alpha}; q, \mathbf{a})|^{2s} \ll (|f(\boldsymbol{\alpha}; X)|^2 - |V(\boldsymbol{\alpha}; q, \mathbf{a})|^2) X^{2s-2} \\ \ll L^2 X^{2s-1}.$$

Since $\mathrm{mes}(\mathfrak{M}) \ll L^{2k+1} X^{-k(k+1)/2}$, we deduce that

$$\int_{\mathfrak{M}} |f(\boldsymbol{\alpha}; X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} - \int_{\mathfrak{M}} |V(\boldsymbol{\alpha})|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \ll (L^{2k+3}/X) X^{2s-k(k+1)/2} \\ = o(X^{2s-k(k+1)/2}). \tag{12.9}$$

By applying [2, Theorems 1.3 and 2.4] as in the argument concluding the proof of [59, Lemma 7.1], one sees that

$$\int_{\mathfrak{M}} |V(\boldsymbol{\alpha})|^{2s} \, \mathrm{d}\boldsymbol{\alpha} = \mathfrak{S}\mathfrak{J},$$

where, when $2s \geqslant \frac{1}{2}k(k+1) + 1$, one has

$$\mathfrak{J} = X^{2s-k(k+1)/2} \int_{\mathbb{R}^k} |I(\boldsymbol{\beta}; 1)|^{2s} \, \mathrm{d}\boldsymbol{\beta} + o(X^{2s-k(k+1)/2}), \tag{12.10}$$

and, when $2s \geqslant \frac{1}{2}k(k+1) + 2$, one has

$$\mathfrak{S} = \sum_{q=1}^{\infty} \sum_{\substack{1 \leqslant \mathbf{a} \leqslant q \\ (q, a_1, \dots, a_k) = 1}} \big| q^{-1} S(q, \mathbf{a}) \big|^{2s} + o(1). \tag{12.11}$$

Here, the integral on the right hand side of (12.10) is absolutely convergent, and the sum on the right hand side of (12.11) is also absolutely convergent. Hence, there exists a real number $C_{s,k} \geqslant 0$ for which

$$\int_{\mathfrak{M}} |V(\boldsymbol{\alpha})|^{2s} \, \mathrm{d}\boldsymbol{\alpha} = C_{s,k} X^{2s-k(k+1)/2} + o(X^{2s-k(k+1)/2}),$$

and we conclude from (12.8) and (12.9) that

$$\oint |f(\boldsymbol{\alpha};X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha} = \int_{\mathfrak{M}} |f(\boldsymbol{\alpha};X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha} + \int_{\mathfrak{m}} |f(\boldsymbol{\alpha};X)|^{2s}\,\mathrm{d}\boldsymbol{\alpha}$$
$$= C_{s,k}X^{2s-k(k+1)/2} + o(X^{2s-k(k+1)/2}). \tag{12.12}$$

Making use of the familiar lower bound $J_{s,k}(X) \gg X^{2s-k(k+1)/2}$, we conclude that when $s > k(k+1)/2$, the asymptotic formula (12.12) holds for some $C_{s,k} > 0$. This completes the proof of Corollary 1.3. □

13. A remark on Tarry's problem

The resolution of the main conjecture in Vinogradov's mean value theorem provides a means of delivering a definitive result concerning Tarry's problem. When h, k and s are positive integers with $h \geqslant 2$, consider the Diophantine system

$$\sum_{i=1}^{s} x_{i1}^j = \sum_{i=1}^{s} x_{i2}^j = \ldots = \sum_{i=1}^{s} x_{ih}^j \quad (1 \leqslant j \leqslant k). \tag{13.1}$$

Let $W(k,h)$ denote the least natural number s having the property that the simultaneous equations (13.1) possess an integral solution $\mathbf{x}$ with

$$\sum_{i=1}^{s} x_{iu}^{k+1} \neq \sum_{i=1}^{s} x_{iv}^{k+1} \quad (1 \leqslant u < v \leqslant h).$$

The problem of estimating $W(k,h)$ has been the subject of extensive investigation by E. M. Wright and L.-K. Hua (see [20], [21] and [60]). There are numerous applications that need not detain us here. However, we note in particular that Croot and Hart [12] have found application of these ideas in work on the sum-product conjecture. Classically, Hua [21] was able to show that $W(k,h) \leqslant k^2(\log k + O(1))$. In recent work [58, Theorem 12.1] based on efficient congruencing, the author was able to improve this conclusion, showing that $W(k,h) \leqslant \frac{1}{2}k(k+1)+1$ for k sufficiently large. We now show that the latter hypothesis on k may be dropped.

Theorem 13.1. *When h and k are natural numbers with $h \geqslant 2$, one has $W(k,h) \leqslant \frac{1}{2}k(k+1)+1$.*

Proof. The argument of the proof of [47, Theorem 1.3] shows that $W(k,h) \leqslant s$ whenever one can establish the bound

$$J_{s,k+1}(X) = o(X^{2s-k(k+1)/2}).$$

However, as a consequence of Corollary 1.3, for all natural numbers k one has $J_{s,k+1}(X) \ll X^{s+\varepsilon}$ whenever $1 \leqslant s \leqslant (k+1)(k+2)/2$. Thus, with $s = \frac{1}{2}k(k+1)+1$, one finds that

$$J_{s,k+1}(X) \ll X^{\varepsilon+1+k(k+1)/2} = X^{\varepsilon-1} \cdot X^{2s-k(k+1)/2}.$$

We therefore conclude that $W(k,h) \leqslant \frac{1}{2}k(k+1)+1$, and the proof of the theorem is complete. □

The bound obtained in Theorem 13.1 apparently achieves the limits of this kind of analytic argument. The best available lower bound for $W(k,h)$ is the trivial bound $W(k,h) \geqslant k+1$, one that for large values of k seems unlikely to represent the true state of affairs. For small values of k, however, explicit numerical examples show that $W(k,2) = k+1$ for $2 \leqslant k \leqslant 9$ and $k = 11$ (see http://euler.free.fr/eslp/eslp.htm).

14. Analogues of Hua's lemma, and Waring's problem

Estimates of the shape (1.3) in Theorem 1.1, and Corollary 1.2, although exhibiting diagonal behaviour, in general fail to provide sufficiently strong mean value estimates that they may be applied directly in Diophantine applications such as Waring's problem. The idea of augmenting the underlying systems with additional low degree equations offers a means of bounding larger moments at a modest cost. As far as the author is aware, this idea seems to date from at least as far back as the work of Arkhipov and Karatsuba from the 1970's, although we have been unable to identify a suitable reference in the literature. In the context of recent advances made via efficient congruencing, this idea has also been noted in conference talks (such as the author's Turán Conference talk in 2011). Most recently, this idea played a pivotal role in the discussion of [6, Theorem 10] associated with an analogue of Hua's lemma. Bourgain reports that an analogue of the proof of the main conjecture in Vinogradov's mean value theorem may be applied to confirm the special case of Corollary 1.2 above in which $(d_1, \dots, d_k) = (1, 2, \dots, k-1, d)$ (see especially [6, equation (6.5)]). In this way, he obtains the bound

$$\int_0^1 \Big| \sum_{1 \leqslant x \leqslant X} e(\alpha x^d) \Big|^{r(r+1)} \mathrm{d}\alpha \ll X^{r^2+\varepsilon} \quad (1 \leqslant r \leqslant d), \tag{14.1}$$

which may be regarded as an analogue of Hua's lemma (see [19]). In this section we apply related ideas to obtain estimates similar in shape to those of estimate (1.3) of Theorem 1.1 and Corollary 1.2, though for higher moments potentially of use in Diophantine applications. These results not only generalise those of Bourgain [6, Theorem 10] to systems of equations, but also generalise them in the case of a single equation in addition to describing the details suppressed in the former treatment (i.e. the proof of Corollary 1.2 in the special case mentioned above).

Throughout this section, we suppose that $\varphi_j \in \mathbb{Z}[t]$ $(1 \leqslant j \leqslant k)$ is a system of polynomials with $\deg(\varphi_j) = d_j$ satisfying the condition

$$1 \leqslant d_k < d_{k-1} < \ldots < d_1. \tag{14.2}$$

It is convenient to adopt the convention that $d_0 = +\infty$ and $d_{k+1} = 0$. We put

$$D = d_1 + \ldots + d_k,$$

and when $r \in \mathbb{N}$, we define

$$\Delta_{r,\mathbf{d}} = \sum_{i=1}^{k} \max\{0, d_i - (r-i+1)\}. \tag{14.3}$$

Finally, we define the exponential sum $F(\boldsymbol{\alpha}) = F(\boldsymbol{\alpha}; \boldsymbol{\varphi})$ by

$$F(\boldsymbol{\alpha}; \boldsymbol{\varphi}) = \sum_{1 \leqslant n \leqslant X} e(\psi(n; \boldsymbol{\alpha})),$$

where $\psi(n; \boldsymbol{\alpha}) = \alpha_1 \varphi_1(n) + \ldots + \alpha_k \varphi_k(n)$.

Theorem 14.1. *For each $r \in \mathbb{N}$ and $\varepsilon > 0$, one has*

$$\int_{[0,1)^k} |F(\boldsymbol{\alpha}; \boldsymbol{\varphi})|^{r(r+1)} \, \mathrm{d}\boldsymbol{\alpha} \ll X^{\varepsilon} \left(X^{r(r+1)/2} + X^{r(r+1)-D+\Delta_{r,\mathbf{d}}} \right).$$

We briefly extract a couple of corollaries.

Corollary 14.2. *Suppose that $\varphi \in \mathbb{Z}[t]$ is a polynomial of degree d. Then for each natural number r with $1 \leqslant r \leqslant d$, and for each $\varepsilon > 0$, one has*

$$\int_0^1 \left| \sum_{1 \leqslant n \leqslant X} e(\alpha \varphi(n)) \right|^{r(r+1)} \mathrm{d}\boldsymbol{\alpha} \ll X^{r^2+\varepsilon}. \tag{14.4}$$

In the case $\varphi(t) = t^d$, this conclusion is just the bound (14.1) described by Bourgain [6, Theorem 10]. The earlier work of Arkhipov and Karatsuba mentioned above would deliver a similar conclusion with the exponent $r(r+1)$ on the left hand side of (14.4) replaced by an integer $s_0 \sim 4r^2 \log r$, and the exponent r^2 on the right hand side replaced by $s_0 - r$. In general, this bound may be applied as a substitute for Hua's lemma [19], which shows under the same hypotheses as in Corollary 14.2 that

$$\int_0^1 \left| \sum_{1 \leqslant n \leqslant X} e(\alpha \varphi(n)) \right|^{2^r} \mathrm{d}\alpha \ll X^{2^r - r + \varepsilon}.$$

Corollary 14.3. *For each $r \in \mathbb{N}$ and $\varepsilon > 0$, one has*

$$\int_{[0,1)^k} |F(\boldsymbol{\alpha}; \boldsymbol{\varphi})|^{r(r+1)} \, \mathrm{d}\boldsymbol{\alpha} \ll X^{r(r+1)/2+\varepsilon}, \quad \text{when } 1 \leqslant r \leqslant k,$$

and when $1 \leqslant m \leqslant k$ and $d_{m+1} + m \leqslant r \leqslant d_m + m - 1$, one has

$$\int_{[0,1)^k} |F(\boldsymbol{\alpha}; \boldsymbol{\varphi})|^{r(r+1)} \, \mathrm{d}\boldsymbol{\alpha} \ll X^{r^2 - (m-1)(2r-m)/2 - d_{m+1} - \ldots - d_k + \varepsilon}.$$

Proof. The first conclusion is immediate from Theorem 14.1, or indeed Theorem 1.1. As for the second, suppose that $1 \leqslant m \leqslant k$ and

$$d_{m+1} + m \leqslant r \leqslant d_m + m - 1.$$

Then one sees that the summands in (14.3) contribute if and only if $1 \leqslant i \leqslant m$, and thus

$$\begin{aligned}\Delta_{r,\mathbf{d}} &= \sum_{i=1}^{m}(d_i - (r-i+1)) \\ &= (D - d_{m+1} - \dots - d_k) - mr + m(m-1)/2.\end{aligned}$$

Hence, we have

$$r(r+1) - D + \Delta_{r,\mathbf{d}} = r^2 - (m-1)r + m(m-1)/2 - d_{m+1} - \dots - d_k,$$

and the desired conclusion follows from Theorem 14.1. □

The proof of Theorem 14.1. When $1 \leqslant r \leqslant k$, the conclusion of Theorem 14.1 is immediate from the case $s = r(r+1)/2$ of the bound (1.3) of Theorem 1.1. We may therefore suppose that $r > k$. The hypothesis (14.2) implies that

$$(d_{i+1} - (r-i)) - (d_i - (r-i+1)) = d_{i+1} - d_i + 1 \leqslant 0,$$

whence

$$d_{i+1} - (r-i) \leqslant d_i - (r-i+1) \quad (0 \leqslant i \leqslant k).$$

Since $d_0 - (r+1) > 0$ and $d_{k+1} - (r-k) < 0$, it follows that there exists an integer l with $0 \leqslant l \leqslant k$ for which

$$d_{l+1} - (r-l) \leqslant 0 \quad \text{and} \quad d_l - (r-l+1) \geqslant 0.$$

We fix any integer l with this property. One then has

$$d_i \leqslant r - i + 1 \quad (l+1 \leqslant i \leqslant k). \tag{14.5}$$

Let $e_1, \dots, e_{r-k}$ denote the distinct positive integers for which

$$\{e_1, \dots, e_{r-k}\} = \{1, 2, \dots, r-l\} \setminus \{d_{l+1}, \dots, d_k\}.$$

Notice that the condition (14.5) ensures that there are indeed $r-k$ such integers. We then define

$$G(\boldsymbol{\alpha}) = \sum_{1 \leqslant n \leqslant X} e(\psi(n;\boldsymbol{\alpha}) + \alpha_{k+1}n^{e_1} + \dots + \alpha_r n^{e_{r-k}}).$$

Finally, for the sake of concision, we write $s = r(r+1)/2$.

By orthogonality, the mean value

$$\oint |F(\boldsymbol{\alpha};\boldsymbol{\varphi})|^{2s}\, \mathrm{d}\boldsymbol{\alpha} \tag{14.6}$$

counts the integral solutions of the system of equations

$$\sum_{i=1}^{s}(\varphi_j(x_i) - \varphi_j(y_i)) = 0 \quad (1 \leqslant j \leqslant k), \tag{14.7}$$

with $1 \leqslant \mathbf{x}, \mathbf{y} \leqslant X$. The mean value (14.6) is therefore equal to the number of integral solutions of the augmented system of equations (14.7) simultaneous with

$$\sum_{i=1}^{s}(x_i^{e_l} - y_i^{e_l}) = h_l \quad (1 \leqslant l \leqslant r-k),$$

with $1 \leqslant \mathbf{x}, \mathbf{y} \leqslant X$ and $|h_l| \leqslant sX^{e_l}$. The point here is that the range for the auxiliary variables h_l is sufficiently large that this new system accommodates all possible choices for $\mathbf{x}$ and $\mathbf{y}$ satisfying (14.7) alone. Thus, by orthogonality and an application of the triangle inequality, we find that the mean value (14.6) is equal to

$$\sum_{|h_1|\leqslant sX^{e_1}} \dots \sum_{|h_{r-k}|\leqslant sX^{e_{r-k}}} \oint |G(\boldsymbol{\beta})|^{2s} e(-\beta_{k+1}h_1 - \dots - \beta_r h_{r-k})\, \mathrm{d}\boldsymbol{\beta} \ll X^{e_1+\dots+e_{r-k}} \oint |G(\boldsymbol{\beta})|^{2s}\, \mathrm{d}\boldsymbol{\beta}.$$

Consequently, one has

$$\oint |F(\boldsymbol{\alpha};\boldsymbol{\varphi})|^{2s}\, \mathrm{d}\boldsymbol{\alpha} \ll X^{(r-l)(r-l+1)/2-d_{l+1}-\dots-d_k} \oint |G(\boldsymbol{\beta})|^{2s}\, \mathrm{d}\boldsymbol{\beta}. \tag{14.8}$$

Next we observe that the Wronskian of the system of polynomials

$$\varphi_1(t), \dots, \varphi_k(t), t^{e_1}, \dots, t^{e_{r-k}} \tag{14.9}$$

may be rearranged so that the polynomials are of increasing degree. The leading monomials are then non-zero integral multiples of

$$t, t^2, \dots, t^{r-l}, t^{d_l}, t^{d_{l-1}}, \dots, t^{d_1}.$$

The Wronskian of the system (14.9) is consequently non-zero, and so it follows from the estimate (1.3) of Theorem 1.1 that for $s = r(r+1)/2$, one has

$$\oint |G(\boldsymbol{\beta})|^{2s}\, \mathrm{d}\boldsymbol{\beta} \ll X^{s+\varepsilon}.$$

On substituting this conclusion into (14.8), we deduce that

$$\oint |F(\boldsymbol{\alpha};\boldsymbol{\varphi})|^{2s}\, \mathrm{d}\boldsymbol{\alpha} \ll X^{\Theta+\varepsilon},$$

where

$$\begin{aligned}\Theta &= \tfrac{1}{2}(r-l)(r-l+1) - d_{l+1} - \dots - d_k + \tfrac{1}{2}r(r+1)\\ &= r(r+1) - lr + \tfrac{1}{2}l(l-1) - D + d_1 + \dots + d_l\\ &= r(r+1) - D + \sum_{i=1}^{l}(d_i - r + i - 1).\end{aligned}$$

Thus we have $\Theta = r(r+1) - D + \Delta_{r,\mathbf{d}}$, where $\Delta_{r,\mathbf{d}}$ is defined via (14.3), and the conclusion of the theorem follows. □

One additional idea in the current repertoire of specialists may, on occasion, offer improvement in the bounds supplied by Theorem 14.1 and its corollaries. In order to describe this idea, we introduce a Hardy-Littlewood dissection. Let $\mathfrak{m} = \mathfrak{m}_\kappa$ denote the set of real numbers $\alpha \in [0,1)$ satisfying the property that, whenever $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $(a,q) = 1$ and $|q\alpha - a| \leqslant (2\kappa)^{-1}X^{1-\kappa}$, then $q > (2\kappa)^{-1}X$. Also, denote by $\mathfrak{M}_\kappa$ the union of the intervals

$$\mathfrak{M}_\kappa(q,a) = \{\alpha \in [0,1) : |q\alpha - a| \leqslant (2\kappa)^{-1}X^{1-\kappa}\},$$

with $0 \leqslant a \leqslant q \leqslant (2\kappa)^{-1}X$ and $(q,a) = 1$. Thus, the unit interval $[0,1)$ is the disjoint union of $\mathfrak{m}_\kappa$ and $\mathfrak{M}_\kappa$. A variant of the proof of [48, Theorem 1.3] yields the following conclusion.

Theorem 14.4. *Suppose that $d_2 \leqslant d_1 - 2$. Then one has the following.*

(i) *When s is a natural number with $2s \geqslant d_1(d_1+1)$, one has*

$$\int_{\mathfrak{m}_{d_1}} \oint |F(\alpha_1, \boldsymbol{\beta}; \boldsymbol{\varphi})|^{2s} \, \mathrm{d}\boldsymbol{\beta} \, \mathrm{d}\alpha_1 \ll X^{2s-D-1+\varepsilon}.$$

(ii) *When s is a natural number with $2s > d_1(d_1-1)$, one has*

$$\int_{\mathfrak{M}_{d_1}} \oint |F(\alpha_1, \boldsymbol{\beta}; \boldsymbol{\varphi})|^{2s} \, \mathrm{d}\boldsymbol{\beta} \, \mathrm{d}\alpha_1 \ll X^{2s-D+\varepsilon}.$$

Proof. Both conclusions follow by applying the argument of the proof of [48, Theorem 2.1], mutatis mutandis. We will be concise with the details in order to save space. Initially, we preserve the option of pursuing either case (i) or case (ii) of the theorem. Thus, we consider $\mathfrak{B} \subseteq [0,1)$, and we define the mean value

$$I(\mathfrak{B}) = \int_{\mathfrak{B}} \oint |F(\alpha_1, \boldsymbol{\beta}; \boldsymbol{\varphi})|^{2s} \, \mathrm{d}\boldsymbol{\beta} \, \mathrm{d}\alpha_1. \tag{14.10}$$

For the sake of clarity and concision, we write κ for d_1. The reader should experience no difficulty in following the argument of the proof of [48, Theorem 2.1] as far as [48, equation (12)], obtaining the bound

$$I(\mathfrak{B}) \ll X^{(\kappa-1)(\kappa-2)/2-(D-\kappa)} \int_{\mathfrak{B}} \oint |H(\alpha_1, \boldsymbol{\theta})|^{2s} \, \mathrm{d}\boldsymbol{\theta} \, \mathrm{d}\alpha_1, \tag{14.11}$$

where

$$H(\alpha_1, \boldsymbol{\theta}) = \sum_{1 \leqslant n \leqslant X} e(\alpha_1 \varphi_1(n) + \theta_1 n + \ldots + \theta_{\kappa-2} n^{\kappa-2}).$$

Here, we have taken integral linear combinations of equations underlying the inner integral of (14.11) so as to reduce to the monomials n^j $(1 \leqslant j \leqslant \kappa - 2)$. Such manœuvring also permits us to assume that $\varphi_1(n)$ takes the shape

$$\varphi_1(n) = An^\kappa + Bn^{\kappa-1},$$

for suitable integers A and B with $A > 0$. Thus, as in [48, equation (13)], we discern that

$$\int_{\mathfrak{B}} \oint |H(\alpha_1, \boldsymbol{\theta})|^{2s} \, \mathrm{d}\boldsymbol{\theta} \, \mathrm{d}\alpha_1 = \sum_{|u| \leqslant sX^{\kappa-1}} \int_{\mathfrak{B}} \oint |h(\alpha_1, \boldsymbol{\beta}; X)|^{2s} e(-\beta_{\kappa-1}u) \, \mathrm{d}\boldsymbol{\beta} \, \mathrm{d}\alpha_1,$$

where

$$h(\alpha_1, \boldsymbol{\beta}; X) = \sum_{1 \leqslant n \leqslant X} e(\psi(n; \alpha_1, \boldsymbol{\beta})),$$

and

$$\psi(n; \alpha_1, \boldsymbol{\beta}) = \alpha_1 \varphi_1(n) + \beta_1 n + \ldots + \beta_{\kappa-1} n^{\kappa-1}.$$

Write

$$K(\gamma) = \sum_{1 \leqslant z \leqslant X} e(-\gamma z) \quad \text{and} \quad \widetilde{K}(\boldsymbol{\gamma}) = \prod_{i=1}^{s} K(\gamma_i) K(-\gamma_{s+i}).$$

In addition, put

$$\mathfrak{h}_y(\alpha_1, \boldsymbol{\beta}; \gamma) = \sum_{1 \leqslant x \leqslant 2X} e(\psi(x-y; \alpha_1, \boldsymbol{\beta}) + \gamma(x-y))$$

and

$$\mathfrak{H}_y(\alpha_1, \boldsymbol{\beta}; \boldsymbol{\gamma}) = \prod_{i=1}^{s} \mathfrak{h}_y(\alpha_1, \boldsymbol{\beta}; \gamma_i) \mathfrak{h}_y(-\alpha_1, -\boldsymbol{\beta}; -\gamma_{s+i}). \tag{14.12}$$

Then, just as in the argument leading to [48, equation (18)], one finds that

$$\int_{\mathfrak{B}} \oint |H(\alpha_1, \boldsymbol{\theta})|^{2s} \, \mathrm{d}\boldsymbol{\theta} \, \mathrm{d}\alpha_1 = \sum_{|u| \leqslant sX^{\kappa-1}} \oint I_u(\boldsymbol{\gamma}, y) \widetilde{K}(\boldsymbol{\gamma}) \, \mathrm{d}\boldsymbol{\gamma}, \tag{14.13}$$

where

$$I_u(\boldsymbol{\gamma}, y) = \int_{\mathfrak{B}} \oint \mathfrak{H}_y(\alpha_1, \boldsymbol{\beta}; \boldsymbol{\gamma}) e(-\beta_{\kappa-1} u) \, \mathrm{d}\boldsymbol{\beta} \, \mathrm{d}\alpha_1.$$

On noting the correction made in [53], the argument leading to [48, equation (22)] yields the bound

$$\sum_{|u| \leqslant sX^{\kappa-1}} I_u(\boldsymbol{\gamma}, y) \ll \int_{\mathfrak{B}} \oint |\mathfrak{H}_0(\alpha_1, \boldsymbol{\beta}; \boldsymbol{\gamma})| \Psi_y(\alpha_1, \beta_{\kappa-1}) \, \mathrm{d}\boldsymbol{\beta} \, \mathrm{d}\alpha_1,$$

where

$$\Psi_y(\alpha_1, \beta_{\kappa-1}) = \Big| \sum_{|u| \leqslant sX^{\kappa-1}} e(-(\kappa A y - B) u \alpha_1 - u \beta_{\kappa-1}) \Big|.$$

Thus we deduce that

$$X^{-1} \sum_{1 \leqslant y \leqslant X} \sum_{|u| \leqslant sX^{\kappa-1}} I_u(\boldsymbol{\gamma}, y) \ll \int_{\mathfrak{B}} \oint |\mathfrak{H}_0(\alpha_1, \boldsymbol{\beta}; \boldsymbol{\gamma})| \Psi(\alpha_1, \beta_{\kappa-1}) \, \mathrm{d}\boldsymbol{\beta} \, \mathrm{d}\alpha_1, \tag{14.14}$$

where

$$\begin{aligned} \Psi(\alpha_1, \beta_{\kappa-1}) &= X^{-1} \sum_{1 \leqslant y \leqslant X} \min\{X^{\kappa-1}, \|(\kappa A y - B)\alpha_1 + \beta_{\kappa-1}\|^{-1}\} \\ &\leqslant X^{-1} \sum_{|z| \leqslant \kappa A X + |B|} \min\{X^{\kappa-1}, \|z\alpha_1 + \beta_{\kappa-1}\|^{-1}\}. \end{aligned}$$

Suppose that $\alpha_1 \in \mathbb{R}$, and that $b \in \mathbb{Z}$ and $r \in \mathbb{N}$ satisfy $(b, r) = 1$ and $|\alpha_1 - b/r| \leqslant r^{-2}$. Then, just as in [48, equation (23)], one obtains the estimate

$$\Psi(\alpha_1, \beta_{\kappa-1}) \ll X^{\kappa-1}(X^{-1} + r^{-1} + rX^{-\kappa}) \log(2r). \tag{14.15}$$

It is at this point that our argument diverges according to whether we are in case (i) or case (ii). We first consider case (i), in which case we put $\mathfrak{B} = \mathfrak{m}_\kappa$. Here, by Dirichlet's approximation theorem, given $\alpha_1 \in \mathfrak{m}_\kappa$, one may find $b \in \mathbb{Z}$ and $r \in \mathbb{N}$ with $(b, r) = 1$, $|r\alpha_1 - b| \leqslant (2\kappa)^{-1} X^{1-\kappa}$ and $r \leqslant 2\kappa X^{\kappa-1}$. The

definition of $\mathfrak{m}_\kappa$ ensures that $r > (2\kappa)^{-1}X$, and hence it follows from (14.15) that

$$\Psi(\alpha_1, \beta_{\kappa-1}) \ll X^{\kappa-2} \log X.$$

From here, the argument leading from [48, equation (23)] to the conclusion of the proof of [48, Theorem 2.1] conveys us via (14.13) and (14.14) to the bound

$$\begin{aligned}\int_{\mathfrak{m}_\kappa} \oint |H(\alpha_1, \boldsymbol{\theta})|^{2s} \, \mathrm{d}\boldsymbol{\theta} \, \mathrm{d}\alpha_1 & \\ &\ll X^{\kappa-2}(\log X) \sup_{\gamma \in [0,1)} \oint |\mathfrak{h}_0(\alpha_1, \boldsymbol{\beta}; \gamma)|^{2s} \, \mathrm{d}\boldsymbol{\beta} \, \mathrm{d}\alpha_1 \oint |\widetilde{K}(\boldsymbol{\gamma})| \, \mathrm{d}\boldsymbol{\gamma} \\ &\ll X^{\kappa-2}(\log X)^{2s+1} J_{s,\kappa}(2X).\end{aligned}$$

When $2s \geqslant \kappa(\kappa+1) = d_1(d_1+1)$, we find from Corollary 1.3 that

$$J_{s,\kappa}(2X) \ll X^{2s-\kappa(\kappa+1)/2+\varepsilon},$$

and thus

$$\int_{\mathfrak{m}_\kappa} \oint |H(\alpha_1, \boldsymbol{\theta})|^{2s} \, \mathrm{d}\boldsymbol{\theta} \, \mathrm{d}\alpha_1 \ll X^{2s-\kappa(\kappa-1)/2-2+\varepsilon}.$$

We therefore conclude from (14.11) that

$$I(\mathfrak{m}_\kappa) \ll X^{2s-D-1+\varepsilon}.$$

In view of the definition (14.10), the first case of the theorem now follows.

Our starting point for the proof of case (ii) of the theorem is again the upper bound (14.15). By appealing to a standard transference principle (see [52, Lemma 14.1]), one deduces that whenever $b \in \mathbb{Z}$ and $r \in \mathbb{N}$ satisfy $(b, r) = 1$ and $|\alpha_1 - b/r| \leqslant r^{-2}$, then one has

$$\Psi(\alpha_1, \beta_{\kappa-1}) \ll X^{\kappa-1+\varepsilon}(\lambda^{-1} + X^{-1} + \lambda X^{-\kappa}),$$

where $\lambda = r + X^\kappa |r\alpha_1 - b|$. When $\alpha \in \mathfrak{M}_\kappa(r, b) \subseteq \mathfrak{M}_\kappa$, moreover, one has $r \leqslant X$ and $X^\kappa |r\alpha_1 - b| \leqslant X$, so that $\lambda \leqslant 2X$. We therefore see that, under such circumstances, one has

$$\Psi(\alpha_1, \beta_{\kappa-1}) \ll X^{\kappa-1+\varepsilon} \Phi(\alpha_1),$$

where $\Phi(\alpha_1)$ is the function taking the value $(q + X^\kappa |q\alpha_1 - a|)^{-1}$, when one has $\alpha_1 \in \mathfrak{M}_\kappa(q, a) \subseteq \mathfrak{M}_\kappa$, and otherwise $\Phi(\alpha_1) = 0$.

It follows from the above discussion that

$$\begin{aligned}\int_{\mathfrak{M}_\kappa} \oint |\mathfrak{h}_0(\alpha_1, \boldsymbol{\beta}; \gamma)|^{2s} \Psi(\alpha_1, \beta_{\kappa-1}) \, \mathrm{d}\boldsymbol{\beta} \, \mathrm{d}\alpha_1 & \\ &\ll X^{\kappa-1+\varepsilon} \int_{\mathfrak{M}_\kappa} \Phi(\alpha_1) \oint |\mathfrak{h}_0(\alpha_1, \boldsymbol{\beta}; \gamma)|^{2s} \, \mathrm{d}\boldsymbol{\beta} \, \mathrm{d}\alpha_1 \\ &\ll X^{\kappa-1+\varepsilon} \int_{\mathfrak{M}_\kappa} \Phi(\alpha_1) \oint |\mathfrak{h}_0(\alpha_1, \boldsymbol{\beta}; 0)|^{2s} \, \mathrm{d}\boldsymbol{\beta} \, \mathrm{d}\alpha_1. \quad (14.16)\end{aligned}$$

Moreover, as a consequence of [10, Lemma 2], we find that

$$\int_{\mathfrak{M}_\kappa} \Phi(\alpha_1) \oint |\mathfrak{h}_0(\alpha_1, \boldsymbol{\beta}; 0)|^{2s} \, \mathrm{d}\boldsymbol{\beta} \, \mathrm{d}\alpha_1 \ll X^{\varepsilon-\kappa}(XI_1 + I_2), \quad (14.17)$$

where

$$I_1 = \int_0^1 \oint |\mathfrak{h}_0(\alpha_1, \boldsymbol{\beta}; 0)|^{2s} \, \mathrm{d}\boldsymbol{\beta} \, \mathrm{d}\alpha_1$$

and

$$I_2 = \oint |\mathfrak{h}_0(0, \boldsymbol{\beta}; 0)|^{2s} \, \mathrm{d}\boldsymbol{\beta}.$$

By appealing to Corollary 1.3, one finds that whenever $2s \geqslant \kappa(\kappa-1)$, one has

$$I_1 \ll J_{s,\kappa}(2X) \ll X^{s+\varepsilon} + X^{2s-\kappa(\kappa+1)/2+\varepsilon}.$$

On the other hand, when $2s \geqslant \kappa(\kappa-1)$, it also follows from Corollary 1.3 that

$$I_2 \ll J_{s,\kappa-1}(2X) \ll X^{2s-\kappa(\kappa-1)/2+\varepsilon}.$$

Provided that $2s \geqslant \kappa(\kappa-1)+2$, therefore, we deduce from (14.17) that

$$\int_{\mathfrak{M}_\kappa} \Phi(\alpha_1) \oint |\mathfrak{h}_0(\alpha_1, \boldsymbol{\beta}; 0)|^{2s} \, \mathrm{d}\boldsymbol{\beta} \, \mathrm{d}\alpha_1 \ll X^{2s-\kappa(\kappa+1)/2+\varepsilon},$$

whence (14.16) yields the estimate

$$\int_{\mathfrak{M}_\kappa} \oint |\mathfrak{h}_0(\alpha_1, \boldsymbol{\beta}; \gamma)|^{2s} \Psi(\alpha_1, \beta_{\kappa-1}) \, \mathrm{d}\boldsymbol{\beta} \, \mathrm{d}\alpha_1 \ll X^{2s-\kappa(\kappa-1)/2-1+\varepsilon}. \tag{14.18}$$

On recalling (14.12), we deduce from (14.18) via Hölder's inequality that

$$\int_{\mathfrak{M}_\kappa} \oint |\mathfrak{H}_0(\alpha_1, \boldsymbol{\beta}; \boldsymbol{\gamma})| \Psi(\alpha_1, \beta_{\kappa-1}) \, \mathrm{d}\boldsymbol{\beta} \, \mathrm{d}\alpha_1 \ll X^{2s-\kappa(\kappa-1)/2-1+\varepsilon},$$

and so (14.14) yields the bound

$$X^{-1} \sum_{1 \leqslant y \leqslant X} \sum_{|u| \leqslant sX^{\kappa-1}} I_u(\boldsymbol{\gamma}, y) \ll X^{2s-\kappa(\kappa-1)/2-1+\varepsilon}.$$

Consequently, much as in the treatment of the previous case, we deduce from (14.13) that

$$\begin{aligned} \int_{\mathfrak{M}_\kappa} \oint |H(\alpha_1, \boldsymbol{\theta})|^{2s} \, \mathrm{d}\boldsymbol{\theta} \, \mathrm{d}\alpha_1 &\ll X^{2s-\kappa(\kappa-1)/2-1+\varepsilon} \oint |\widetilde{K}(\boldsymbol{\gamma})| \, \mathrm{d}\boldsymbol{\gamma} \\ &\ll X^{2s-\kappa(\kappa-1)/2-1+2\varepsilon}. \end{aligned}$$

On substituting this estimate into (14.11), we find that

$$I(\mathfrak{M}_\kappa) \ll X^{2s-D+2\varepsilon}.$$

In view of the definition (14.10), the conclusion of the theorem now follows in case (ii). □

By combining the two conclusions of Theorem 14.4, we obtain a slight improvement on Theorem 14.1 for larger moments.

Theorem 14.5. *Suppose that $d_2 \leqslant d_1 - 2$. Put*

$$u = d_1(d_1+1) - \max_{k \leqslant r \leqslant d_1} \frac{d_1(d_1+1) - r(r+1)}{1 + \Delta_{r,\mathbf{d}}}. \tag{14.19}$$

Then whenever s is a real number with $2s \geqslant \max\{u, d_1(d_1-1)+2\}$, one has

$$\oint |F(\boldsymbol{\alpha};\boldsymbol{\varphi})|^{2s}\,\mathrm{d}\boldsymbol{\alpha} \ll X^{2s-D+\varepsilon}.$$

Proof. Suppose that the maximum in the definition (14.19) occurs for the index r, and put $\Delta = \Delta_{r,\mathbf{d}}$. Then it follows from Hölder's inequality that

$$\int_{\mathfrak{m}_{d_1}} \oint |F(\alpha_1,\boldsymbol{\beta};\boldsymbol{\varphi})|^u\,\mathrm{d}\boldsymbol{\beta}\,\mathrm{d}\alpha_1 \leqslant T_1^{\Delta/(1+\Delta)} T_2^{1/(1+\Delta)},$$

where

$$T_1 = \int_{\mathfrak{m}_{d_1}} \oint |F(\alpha_1,\boldsymbol{\beta};\boldsymbol{\varphi})|^{d_1(d_1+1)}\,\mathrm{d}\boldsymbol{\beta}\,\mathrm{d}\alpha_1$$

and

$$T_2 = \int_0^1 \oint |F(\alpha_1,\boldsymbol{\beta};\boldsymbol{\varphi})|^{r(r+1)}\,\mathrm{d}\boldsymbol{\beta}\,\mathrm{d}\alpha_1.$$

Thus, from Theorems 14.1 and 14.4(i), one finds that

$$\int_{\mathfrak{m}_{d_1}} \oint |F(\alpha_1,\boldsymbol{\beta};\boldsymbol{\varphi})|^u\,\mathrm{d}\boldsymbol{\beta}\,\mathrm{d}\alpha_1 \ll X^{\varepsilon-D}\left(X^{d_1(d_1+1)-1}\right)^{\Delta/(1+\Delta)}\left(X^{r(r+1)+\Delta}\right)^{1/(1+\Delta)}$$
$$\ll X^{u-D+\varepsilon}.$$

Meanwhile, provided that $u \geqslant d_1(d_1-1)+2$, one finds from Theorem 14.4(ii) that

$$\int_{\mathfrak{M}_{d_1}} \oint |F(\alpha_1,\boldsymbol{\beta};\boldsymbol{\varphi})|^u\,\mathrm{d}\boldsymbol{\beta}\,\mathrm{d}\alpha_1 \ll X^{u-D+\varepsilon}.$$

Combining these two estimates, we see that

$$\oint |F(\boldsymbol{\alpha};\boldsymbol{\varphi})|^u\,\mathrm{d}\boldsymbol{\alpha} \ll X^{u-D+\varepsilon},$$

and the conclusion of the theorem follows. □

Three corollaries of Theorem 14.5 may be of interest. First we consider the mean value

$$I_{s,d}(X) = \int_{[0,1)^{d-1}} \left| \sum_{1\leqslant x\leqslant X} e(\alpha_d x^d + \alpha_{d-2}x^{d-2} + \ldots + \alpha_1 x)\right|^{2s} \mathrm{d}\boldsymbol{\alpha},$$

in which the argument of the exponential sum is a polynomial of degree d in which there is no monomial of degree $d-1$. Hua investigated the problem of determining the smallest positive integer S_d having the property that whenever $2s \geqslant S_d$, then

$$I_{s,d}(X) \ll X^{2s-(d^2-d+2)/2+\varepsilon}. \tag{14.20}$$

Here, since the sum of the degrees in the associated Diophantine system of equations is

$$1+2+\ldots+(d-2)+d = (d^2-d+2)/2,$$

the bound Hua sought is essentially best possible for $s \geqslant (d^2-d+2)/2$. This mean value played a critical role in his approach to Vinogradov's mean

value theorem for small degrees (see [23, Chapter 5]). Thus, Hua obtained the bounds

$$S_3 \leqslant 10, \quad S_4 \leqslant 32, \quad S_5 \leqslant 86, \dots .$$

More recently, as a consequence of progress on Vinogradov's mean value theorem stemming from the efficient differencing method, the author obtained the bounds $S_k \leqslant 2k^2 - 2k$ (see [49, Theorem 11.6]) and $S_3 \leqslant 9$ (see [51, Theorem 1.1]).

As a consequence of Theorem 14.5, we obtain new bounds for S_d for $d \geqslant 4$.

Corollary 14.6. *When $d \geqslant 3$, one has $S_d \leqslant d^2$.*

Proof. Take $d_1 = d$ and $d_i = d - i$ $(2 \leqslant i \leqslant d-1)$, and put $k = d-1$. We apply Theorem 14.5 with $r = d-1$. In such circumstances, we find from (14.3) that

$$\Delta_{r,\mathbf{d}} = d - r + \sum_{i=2}^{d-1} ((d-i) - (r+1-i)) = 1.$$

Thus, on putting

$$u = d(d+1) - \frac{d(d+1) - d(d-1)}{2} = d^2,$$

it follows from Theorem 14.5 that the upper bound (14.20) holds whenever $2s \geqslant u$. Thus $S_d \leqslant u = d^2$, and the proof of the corollary is complete. □

Next we consider Waring's problem. When s and d are natural numbers, let $R_{s,d}(n)$ denote the number of representations of the natural number n as the sum of s dth powers of positive integers. A formal application of the circle method suggests that for $d \geqslant 3$ and $s \geqslant d+1$, one should have

$$R_{s,d}(n) = \frac{\Gamma(1+1/d)^s}{\Gamma(s/d)} \mathfrak{S}_{s,d}(n) n^{s/d-1} + o(n^{s/d-1}), \tag{14.21}$$

where

$$\mathfrak{S}_{s,d}(n) = \sum_{q=1}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \left(q^{-1} \sum_{r=1}^{q} e(ar^d/q) \right)^s e(-na/q).$$

Granted appropriate congruence conditions on n, one has $1 \ll \mathfrak{S}_{s,d}(n) \ll n^{\varepsilon}$, so that the conjectured relation (14.21) is a legitimate asymptotic formula. Let $\widetilde{G}(d)$ denote the least integer t with the property that, for all $s \geqslant t$, and all sufficiently large natural numbers n, one has the asymptotic formula (14.21). We numerically sharpen the conclusion $\widetilde{G}(d) \leqslant d^2 - d + O(\sqrt{d})$ recorded by Bourgain [6, Theorem 11], achieving the limit of the method.

We define the integer $\theta = \theta(d)$ by

$$\theta(d) = \begin{cases} 1, & \text{when } 2d+2 \geqslant \lfloor \sqrt{2d+2} \rfloor^2 + \lfloor \sqrt{2d+2} \rfloor, \\ 2, & \text{when } 2d+2 < \lfloor \sqrt{2d+2} \rfloor^2 + \lfloor \sqrt{2d+2} \rfloor. \end{cases} \tag{14.22}$$

Corollary 14.7. *Let*

$$s_0 = d(d-1) + \min_{0\leqslant m<d} \frac{2d + m(m-1)}{m+1}.$$

Then, whenever $s \geqslant s_0$, one has

$$\int_0^1 \Big| \sum_{1\leqslant n\leqslant X} e(\alpha n^d) \Big|^s \, \mathrm{d}\boldsymbol{\alpha} \ll X^{s-d+\varepsilon}. \tag{14.23}$$

Thus, one has $\widetilde{G}(d) \leqslant \lfloor s_0 \rfloor + 1$, and in particular

$$\widetilde{G}(d) \leqslant d^2 - d + 2\lfloor \sqrt{2d+2} \rfloor - \theta(d).$$

Proof. It is apparent from Corollary 14.2 that $\Delta_{r,d} = d - r$ $(1 \leqslant r \leqslant d)$. Then it follows from Theorem 14.5 that the estimate (14.23) holds whenever $s \geqslant \max\{u, d(d-1)+2\}$, where

$$\begin{aligned} u &= d(d+1) - \max_{0\leqslant m<d} \frac{d(d+1) - (d-m)(d-m+1)}{m+1} \\ &= d(d+1) - \max_{0\leqslant m<d} \frac{2dm - m(m-1)}{m+1} = s_0. \end{aligned}$$

It is apparent that $s_0 \geqslant d(d-1)+2$, and hence the conclusion (14.23) holds whenever $s \geqslant s_0$. Moreover, granted the estimate (14.23) in the case $s = s_0$, the methods of [38, Chapter 4] show that $\widetilde{G}(k) \leqslant \lfloor s_0 \rfloor + 1$.

All that remains to complete the proof of the corollary is the confirmation of the final bound on $\widetilde{G}(d)$, and this we obtain by deriving an explicit bound on s_0. Take $m = \lfloor \sqrt{2d+2} \rfloor$, and define ω via the relation $\sqrt{2d+2} = m + \omega$. Then we have $0 \leqslant \omega < 1$. With this choice of m, one finds that

$$\begin{aligned} \frac{2d + m(m-1)}{m+1} &= \frac{(m+\omega)^2 - 2 + m(m-1)}{m+1} \\ &= \frac{2m(m+1) - (3-2\omega)(m+1) + (1-\omega)^2}{m+1}, \\ &= 2m - 3 + \delta, \end{aligned}$$

where

$$\delta = 2\omega + \frac{(1-\omega)^2}{m+1}. \tag{14.24}$$

In all circumstances, one has

$$\delta < 2\omega + (1-\omega)^2/2 = (1+\omega)^2/2 < 2,$$

whence

$$\frac{2d + m(m-1)}{m+1} < 2m - 1.$$

Then we may take $s_0 = w$ with $w < d^2 - d + 2\lfloor \sqrt{2d+2} \rfloor - 1$, yielding the bound

$$\widetilde{G}(d) \leqslant \lfloor w \rfloor + 1 \leqslant d^2 - d + 2\lfloor \sqrt{2d+2} \rfloor - 1.$$

On the other hand, provided that $\delta < 1$, we instead obtain

$$\frac{2d + m(m-1)}{m+1} < 2m - 2.$$

In such circumstances, we may take $s_0 = w$ with $w < d^2 - d + 2\lfloor\sqrt{2d+2}\rfloor - 2$, delivering the bound

$$\widetilde{G}(d) \leqslant \lfloor w \rfloor + 1 \leqslant d^2 - d + 2\lfloor\sqrt{2d+2}\rfloor - 2.$$

It follows from (14.24) that $\delta < 1$ if and only if

$$(2m+2)\omega + (1-\omega)^2 < m + 1,$$

or equivalently

$$2d + 2 = (m+\omega)^2 < m^2 + m.$$

This completes the proof of the corollary. □

We remark that it seems that no improvement in the conclusion of Corollary 14.7 is gained by taking $m = \lceil\sqrt{2d+2}\rceil$, so that the stated bounds on $\widetilde{G}(d)$ are the sharpest obtainable using this circle of ideas. The formula for s_0 in the statement of the corollary is equivalent to that given by Bourgain [6, Theorem 11]. We note that, as shown in [48, Theorem 4.1], the truth of the main conjecture in Vinogradov's mean value theorem (Corollary 1.3 or [8]) delivers the bounds $\widetilde{G}(4) \leqslant 15$, $\widetilde{G}(5) \leqslant 23$, $\widetilde{G}(6) \leqslant 34$, $\widetilde{G}(7) \leqslant 47$, $\widetilde{G}(8) \leqslant 61$, $\widetilde{G}(9) \leqslant 78$, $\widetilde{G}(10) \leqslant 97$, and so on. The conclusion of Corollary 14.7 matches or improves on these bounds for $k \geqslant 10$.

We finish by briefly outlining how Theorem 14.5 may be applied to treat systems with one large and a number of smaller degree terms with an efficiency matching Corollary 14.7. We again make use of the definition (14.22) of the integer $\theta(d)$.

Corollary 14.8. *Suppose that*

$$d_2 \leqslant d_1 - \lfloor\sqrt{2d_1+2}\rfloor - 1. \tag{14.25}$$

Then there is a positive number τ having the property that, with

$$s_0 = d_1^2 - d_1 + 2\lfloor\sqrt{2d_1+2}\rfloor - \theta(d_1) - \tau,$$

one has

$$\oint |F(\boldsymbol{\alpha};\boldsymbol{\varphi})|^{s_0}\,\mathrm{d}\boldsymbol{\alpha} \ll X^{s_0-D}.$$

Proof. Write $m = \lfloor\sqrt{2d_1+2}\rfloor$. Under the hypothesis (14.25), it is apparent that $d_1 \geqslant d_2 + 2$. Also, on taking $r = d_1 - m$, we find that $d_2 - r + 1 \leqslant 0$. We therefore deduce from (14.3) that $\Delta_{r,\mathbf{d}} = d_1 - r = m$. Thus, just as in the proof of Corollary 14.7, if we put

$$u = d_1(d_1+1) - \frac{d_1(d_1+1) - (d_1-m)(d_1-m+1)}{m+1},$$

then we find that $u < d_1(d_1-1)+2m-\theta$. Since $u \geqslant d_1(d_1-1)+2$, then again as in the proof of Corollary 14.7, we see that when s is a real number with $2s \geqslant u$, one has

$$\oint |F(\boldsymbol{\alpha};\boldsymbol{\varphi})|^{2s}\,\mathrm{d}\boldsymbol{\alpha} \ll X^{2s-D+\varepsilon}.$$

In order to complete the proof of the corollary, we have now only to apply the Hardy-Littlewood method. The details are standard, and so we offer only the briefest outline of the necessary argument. Take u_0 to be a real number satisfying

$$u < u_0 < d_1(d_1-1)+2\lfloor\sqrt{2d_1+2}\rfloor-\theta,$$

and put

$$\tau = d_1(d_1-1)+2\lfloor\sqrt{2d_1+2}\rfloor-\theta-u_0.$$

We then take $\delta = \tau/(100d_1)$. We define the set of major arcs $\mathfrak{N}$ to be the union of the arcs

$$\mathfrak{N}(q,\mathbf{a}) = \{\boldsymbol{\alpha}\in[0,1)^k : |\alpha_i - a_i/q| \leqslant X^{\delta-d_i}\ (1\leqslant i\leqslant k)\},$$

with

$$0\leqslant \mathbf{a}\leqslant q,\quad q\leqslant X^{\delta}\quad\text{and}\quad (q,a_1,\dots,a_k)=1.$$

Also, we put $\mathfrak{n} = [0,1)^k\setminus\mathfrak{N}$. Then it follows from [47, Theorem 1.6] that whenever $|F(\boldsymbol{\alpha};\boldsymbol{\varphi})| > X^{1-\delta/d_1^3}$, then $\boldsymbol{\alpha}\in\mathfrak{N}$. Thus

$$\begin{aligned}\int_{\mathfrak{n}} |F(\boldsymbol{\alpha};\boldsymbol{\varphi})|^{u_0+\tau}\,\mathrm{d}\boldsymbol{\alpha} &\ll (X^{1-\delta/d_1^3})^{\tau}\oint |F(\boldsymbol{\alpha};\boldsymbol{\varphi})|^{u_0}\,\mathrm{d}\boldsymbol{\alpha}\\ &\ll X^{u_0+\tau-D}.\end{aligned}$$

Meanwhile, by applying the methods based on [2, Theorems 1.3 and 2.4], just as in the proof of Corollary 1.3, one obtains the bound

$$\int_{\mathfrak{N}} |F(\boldsymbol{\alpha};\boldsymbol{\varphi})|^{u_0+\tau} \ll X^{u_0+\tau-D}.$$

By combining these estimates, the conclusion of the corollary follows. □

When $k=2$, $d_1=d$ and $d_2=1$, the conclusion of Corollary 14.8 shows that the estimate

$$\int_{[0,1)^2}\Big|\sum_{1\leqslant n\leqslant X} e(\alpha_1 n^d+\alpha_2 n)\Big|^s\,\mathrm{d}\boldsymbol{\alpha} \ll X^{s-d-1} \tag{14.26}$$

holds whenever $s>u_0$, for some real number u_0 with

$$u_0 < d(d-1)+2\lfloor\sqrt{2d+2}\rfloor-\theta(d),$$

provided at least that one has $d-\lfloor\sqrt{2d+2}\rfloor\geqslant 2$. This condition is satisfied for $d\geqslant 5$, as is readily confirmed. For small values of d, the methods of Hua [23] play a role (see also [11, Lemma 5]), for one has the bound

$$\int_{[0,1)^2}\Big|\sum_{1\leqslant n\leqslant X} e(\alpha_1 n^d+\alpha_2 n)\Big|^{2^j+2}\,\mathrm{d}\boldsymbol{\alpha} \ll X^{2^j-j+1+\varepsilon}\quad (2\leqslant j\leqslant d). \tag{14.27}$$

By applying this estimate as a substitute for Theorem 14.1 in the proof of Theorem 14.5, we find by applying Hölder's inequality that the estimate (14.26) holds for $s > s_0(d)$, where

$$s_0(4) = 15, \quad s_0(5) = 23\tfrac{1}{3}, \quad s_0(6) = 34, \quad s_0(7) = 46\tfrac{1}{2},$$
$$s_0(8) = 61\tfrac{1}{5}, \quad s_0(9) = 78, \quad s_0(10) = 96\tfrac{6}{7}.$$

Here, one makes use of the case $j = 3$ of the bound (14.27) for $d \leqslant 6$, and $j = 4$ for $7 \leqslant d \leqslant 10$. Meanwhile, the work of [51, Theorem 1.1] shows that when $d = 3$, then (14.26) holds whenever $s > 9$. We remark that, in this special case $k = 2$, $d_1 = d$ and $d_2 = 1$, very slightly weaker bounds could be extracted from Table 1 of the paper [1] that was submitted to the arXiv just prior to the submission of this memoir. The underlying minor arc bounds can be seen to be morally equivalent, though the major arc treatment differs.

15. Vinogradov's mean value theorem in number fields

The application of the Hardy-Littlewood (circle) method in number fields is frequently complicated by the dependence of exponential sum estimates on the degree of the ambient field extension. In Diophantine problems of all but the lowest degrees d, existing methods for circumventing such difficulties demand the availability of a number of variables exponentially large in terms of d. Thus, for example, Birch [4] has shown that in any algebraic number field, the rational solutions of a diagonal form of degree d in s variables have the expected asymptotic density whenever $s \geqslant 2^d + 1$. The approach of Birch owes its success to the efficiency of Hua's lemma with 2^d variables. Indeed, so efficient is the latter that, equipped with even a weak version of Weyl's inequality having poor dependence on the degree of the field extension at hand, a satisfactory outcome can be derived with just one additional variable. Hitherto, such efficiency has been absent from versions of Vinogradov's mean value theorem that might otherwise be expected to deliver superior bounds for the number of variables (see, for example, the work of Körner [24] and Eda [13]). Our primary goal in this section is to establish such an efficient version of Vinogradov's mean value theorem in number fields, thereby opening access to sharp Diophantine applications in number fields. Indeed, we establish the main conjecture in number fields, and this delivers an analogue of Birch's theorem whenever $s \geqslant d^2 + d + 1$.

In order to be more concrete concerning our conclusions, we require some notation, beginning with the infrastructure for algebraic number fields. We refer the reader to [42] for an introduction to the circle method in number fields. We consider an algebraic extension K of degree n over $\mathbb{Q}$. Let $K^{(l)}$ $(1 \leqslant l \leqslant n_1)$ be the real conjugate fields associated with K, and let $K^{(m)}$ and $K^{(m+n_2)}$ $(n_1 + 1 \leqslant m \leqslant n_1 + n_2)$ be the pairs of complex conjugate fields associated with K. Here, one has $n_1 + 2n_2 = n$. We write $\mathfrak{O}_K$ for the ring of integers of K, and we fix a basis $\Omega = \{\omega_1, \dots, \omega_n\}$ for $\mathfrak{O}_K$ over $\mathbb{Z}$. We then denote by $\mathcal{B}(X) \subset \mathfrak{O}_K$ the unit cube

$$\mathcal{B}(X) = \big\{r_1\omega_1 + \ldots + r_n\omega_n : r_i \in [-\tfrac{1}{2}X^{1/n}, \tfrac{1}{2}X^{1/n}) \cap \mathbb{Z} \ (1 \leqslant i \leqslant n)\big\}.$$

It is apparent that $\mathrm{card}(\mathcal{B}(X)) \asymp X$.

When $\gamma \in K$, we denote by $\gamma^{(i)}$ the conjugates of γ, where $\gamma^{(i)} \in K^{(i)}$ $(1 \leqslant i \leqslant n)$. Then, as usual, we define the trace map $\mathrm{Tr} = \mathrm{Tr}_{K/\mathbb{Q}}$ and norm map $\mathrm{N} = \mathrm{N}_{K/\mathbb{Q}}$ by taking

$$\mathrm{Tr}(\gamma) = \gamma^{(1)} + \ldots + \gamma^{(n)} \quad \text{and} \quad \mathrm{N}(\gamma) = \gamma^{(1)} \cdots \gamma^{(n)}.$$

When $\gamma_j \in K$ and $\theta_j \in \mathbb{R}$ for $1 \leqslant j \leqslant n$, and

$$\lambda(\boldsymbol{\theta}) = \theta_1\gamma_1 + \ldots + \theta_n\gamma_n,$$

we define

$$\lambda^{(i)}(\boldsymbol{\theta}) = \theta_1\gamma_1^{(i)} + \ldots + \theta_n\gamma_n^{(i)} \quad (1 \leqslant i \leqslant n).$$

The trace map on K can then be extended by defining

$$\mathrm{Tr}(\lambda(\boldsymbol{\theta})) = \lambda^{(1)}(\boldsymbol{\theta}) + \ldots + \lambda^{(n)}(\boldsymbol{\theta}).$$

The analogue of the function $e(\alpha) = e^{2\pi i\alpha}$ in this number field setting is then defined by taking $E(\lambda(\boldsymbol{\theta})) = e(\mathrm{Tr}(\lambda(\boldsymbol{\theta})))$.

Next, let $\mathfrak{d}^{-1}$ denote the inverse different, so that

$$\mathfrak{d}^{-1} = \{\eta \in K : \mathrm{Tr}(\eta\xi) \in \mathbb{Z} \text{ for all } \xi \in \mathfrak{O}_K\}.$$

There is a basis $\mathrm{P} = \{\rho_1, \ldots, \rho_n\}$ of $\mathfrak{d}^{-1}$ dual to Ω having the property that

$$\mathrm{Tr}(\rho_i\omega_j) = \begin{cases} 1, & \text{when } i = j, \\ 0, & \text{when } i \neq j. \end{cases}$$

As a special case of the linear form $\lambda(\boldsymbol{\theta})$ defined above, we define $\alpha = \alpha(\boldsymbol{\theta})$ by

$$\alpha(\boldsymbol{\theta}) = \theta_1\rho_1 + \ldots + \theta_n\rho_n.$$

When such a form occurs in an n-fold integral, we use the symbol $\mathrm{d}\alpha$ to denote the n-fold differential $\mathrm{d}\theta_1 \ldots \mathrm{d}\theta_n$. It is convenient then to write $\mathbb{T}$ for $[0,1)^n$. Now that we are equipped with this notation, we may record the fundamental orthogonality relation that underpins the circle method in number fields. Thus, when $\gamma \in \mathfrak{O}_K$, one has

$$\int_{\mathbb{T}} E(\alpha\gamma)\,\mathrm{d}\alpha = \begin{cases} 1, & \text{when } \gamma = 0, \\ 0, & \text{when } \gamma \in \mathfrak{O}_K \setminus \{0\}. \end{cases} \tag{15.1}$$

We may now announce the analogue of Theorem 1.1 in number fields.

Theorem 15.1. *Suppose that $\varphi_j \in \mathfrak{O}_K[t]$ $(1 \leqslant j \leqslant k)$ is a system of polynomials with $W(t;\boldsymbol{\varphi}) \neq 0$. Let s be a positive real number with $s \leqslant k(k+1)/2$. Also, suppose that $(\mathfrak{a}_\nu)_{\nu\in\mathfrak{O}_K}$ is a sequence of complex numbers. Then for each $\varepsilon > 0$, one has*

$$\int_{\mathbb{T}^k} \Big| \sum_{\nu\in\mathcal{B}(X)} \mathfrak{a}_\nu E(\alpha_1\varphi_1(\nu) + \ldots + \alpha_k\varphi_k(\nu)) \Big|^{2s} \mathrm{d}\boldsymbol{\alpha} \ll X^{\varepsilon} \Big(\sum_{\nu\in\mathcal{B}(X)} |\mathfrak{a}_\nu|^2 \Big)^s.$$

In particular, one has

$$\int_{\mathbb{T}^k}\Big|\sum_{\nu\in\mathcal{B}(X)}E(\alpha_1\varphi_1(\nu)+\ldots+\alpha_k\varphi_k(\nu))\Big|^{2s}\,\mathrm{d}\boldsymbol{\alpha}\ll X^{s+\varepsilon}.$$

Throughout this section, we adopt the convention that implicit constants in Vinogradov's notation $\ll$ and $\gg$ may depend on s, k, $\boldsymbol{\varphi}$, K, Ω, and also the small positive number ε. As an immediate consequence of the orthogonality relation (15.1), one obtains the following corollary.

Corollary 15.2. *When $s\in\mathbb{N}$, denote by $\mathcal{N}_{s,\boldsymbol{\varphi}}(X;K)$ the number of solutions of the system of equations*

$$\sum_{i=1}^{s}(\varphi_j(x_i)-\varphi_j(y_i))=0\quad(1\leqslant j\leqslant k),$$

with $x_i,y_i\in\mathcal{B}(X)$. Then whenever $W(t;\boldsymbol{\varphi})\neq 0$ and $s\leqslant k(k+1)/2$, one has

$$\mathcal{N}_{s,\boldsymbol{\varphi}}(X;K)\ll X^{s+\varepsilon}.$$

Finally, when $k\in\mathbb{N}$ and $s>0$, we define

$$J_{s,k}(X;K)=\int_{\mathbb{T}^k}\Big|\sum_{\nu\in\mathcal{B}(X)}E(\alpha_1\nu+\ldots+\alpha_k\nu^k)\Big|^{2s}\,\mathrm{d}\boldsymbol{\alpha}.$$

Theorem 15.1 delivers an analogue of the main conjecture in Vinogradov's mean value theorem for algebraic number fields.

Corollary 15.3. *Suppose that $k\in\mathbb{N}$ and $s>0$. Then for each $\varepsilon>0$, one has*

$$J_{s,k}(X;K)\ll X^{\varepsilon}(X^s+X^{2s-k(k+1)/2}).$$

The literature concerning Vinogradov's mean value theorem in number fields begins with the work of Körner [24] more than half a century ago. When $[K:\mathbb{Q}]=n$ and $s\geqslant\frac{1}{4}nk(k+1)+rk$ ($r\in\mathbb{N}$), Körner [24, Satz 1] delivers an estimate tantamount to

$$J_{s,k}(X;K)\ll X^{2s-\frac{1}{2}k(k+1)+\eta_{s,k}}(\log X)^r,$$

where $\eta_{s,k}=\frac{1}{2}k(k+1)(1-1/k)^r$. This was improved by Eda [13] (see also [14, Lemma 4]) to the extent that the power of $\log X$ may be deleted, and the condition on s relaxed to

$$s\geqslant\frac{n}{n-1}k(k+1)+rk-1.$$

Recent work of Kozlov [25] and Sorokin [35] provides some slight improvement in these results for the special case $K=\mathbb{Q}(\sqrt{-1})$, although their estimates are constrained to possess the same salient features. Thus, in all of this previous work, the exponent $\eta_{s,k}$ behaves roughly like k^2e^{-s/k^2}. In consequence, one must take s to be at least as large as $k^2(2\log k+\log\log k+c)$, for a suitable positive constant c, in order that the quality of available estimates of Weyl type permit sufficient control of the mean value implicit in $J_{s,k}(X;K)$ necessary for

applications. Indeed, the dependence of available Weyl estimates on the degree n of the field extension may necessitate that this constant c grow with n at least as fast as $\log n$. By contrast, Corollary 15.3 permits full control to be exercised as soon as $s \geqslant k(k+1)/2$. Not only is the dependence on the degree of the ambient field extension entirely removed, but the dependence on k is also substantially improved.

The conclusion of Theorem 15.1 follows from an analogue of Theorem 3.1, as we now describe. Since the details are strikingly similar to those in the situation over $\mathbb{Z}$ described in §§3-12, we will economise on space by indicating only the places in the argument where special care must be taken. We again take k to be an integer with $k \geqslant 1$, and consider polynomials $\varphi_1, \dots, \varphi_k \in \mathfrak{O}_K[t]$. Throughout the argument, the ring of integers $\mathfrak{O}_K$ replaces $\mathbb{Z}$. Let $\mathfrak{p} = (\pi)$ be a prime ideal of $\mathfrak{O}_K$, and put $p = \mathrm{N}(\pi)$. We will be implicitly working in the $\mathfrak{p}$-adic field completing K at the place $\mathfrak{p}$. We assume throughout that $p > (k!)^n$. The definitions of §3 must now be made, mutatis mutandis, where throughout we emphasise that $\mathbb{Z}$ is replaced by $\mathfrak{O}_K$, congruences mod p^h are replaced by congruences mod $\mathfrak{p}^h$, and the function $e(z)$ is replaced by $E(z)$. The definitions (1.9) and (3.5) must be adjusted to this new setting. First, when $F : \mathbb{T}^k \to \mathbb{C}$ is integrable, we write

$$\oint F(\boldsymbol{\alpha})\, \mathrm{d}\boldsymbol{\alpha} = \int_{\mathbb{T}^k} F(\boldsymbol{\alpha})\, \mathrm{d}\boldsymbol{\alpha}.$$

Next, when B is a positive integer, we define

$$\oint_{\mathfrak{p}^B} F(\boldsymbol{\alpha})\, \mathrm{d}\boldsymbol{\alpha} = p^{-kB} \sum_{u_1 \bmod \mathfrak{p}^B} \dots \sum_{u_k \bmod \mathfrak{p}^B} F(\mathbf{u}\pi^{-B}). \tag{15.2}$$

Some words of explanation are in order here. First, the summations on the right hand side of (15.2) are taken over complete sets of residues modulo $\mathfrak{p}^B$. Next, each coordinate $u_i\pi^{-B}$ of the argument on the right hand side of (15.2) may be written in the shape

$$u_i = \beta_{i1}\rho_1 + \dots + \beta_{in}\rho_n,$$

with $\beta_{ij} \in \mathbb{R}$, and we then reduce each coefficient β_{ij} modulo 1. With this convention, we may regard $\mathbf{u}\pi^{-B}$ as belonging to $\mathbb{T}^k$. Integrals with subscripts p^B in §§3-12 are now replaced by this newly defined integral with subscript $\mathfrak{p}^B$, in the obvious fashion. With this definition, one may verify that for $\nu \in \mathfrak{O}_K$, one has the orthogonality relation

$$\oint_{\mathfrak{p}^B} E(\alpha\nu)\, \mathrm{d}\alpha = \begin{cases} 1, & \text{when } \nu \equiv 0 \pmod{\mathfrak{p}^B}, \\ 0, & \text{when } \nu \not\equiv 0 \pmod{\mathfrak{p}^B}. \end{cases}$$

Equipped with these modified definitions, a version of Theorem 3.1 may now be stated in the number field setting. We recall in advance the definitions (3.6) and (3.8) in their modified manifestations.

Theorem 15.4. *Let K be an algebraic extension of $\mathbb{Q}$ with $[K : \mathbb{Q}] < \infty$. Suppose that $k \in \mathbb{N}$, and that $\mathfrak{p} = (\pi)$ is a prime ideal of $\mathfrak{O}_K$ with $p = \mathrm{N}_{K/\mathbb{Q}}(\pi) > (k!)^n$. Then one has $\lambda(k(k+1)/2, k) = 0$.*

We remark that the condition $p > (k!)^n$ is imposed in order to ensure that any factor $k!$, occurring in a binomial expansion en route, is necessarily non-zero modulo $\mathfrak{p}$. If one were to have $k! \equiv 0 \pmod{\mathfrak{p}}$, then one would have that $\mathrm{N}(\pi)$ divides $\mathrm{N}(k!)$ over $\mathbb{Z}$. But by the aforementioned hypothesis, one has $\mathrm{N}(\pi) > (k!)^n \geqslant \mathrm{N}(k!)$, leading to a contradiction, and so the desired conclusion follows.

Corollary 15.5. *Let K be an algebraic extension of $\mathbb{Q}$ with $[K : \mathbb{Q}] < \infty$. Suppose that $k \in \mathbb{N}$, and that $\mathfrak{p} = (\pi)$ is a prime ideal of $\mathfrak{O}_K$ with $p = \mathrm{N}_{K/\mathbb{Q}}(\pi) > (k!)^n$. Suppose that $\tau > 0$ and $\varepsilon > 0$. Let B be sufficiently large in terms of n, k, τ and ε. Put $s = k(k+1)/2$ and $H = \lceil B/k \rceil$. Then for every $\boldsymbol{\varphi} \in \Phi_\tau(B)$, and every sequence $(\mathfrak{a}_\nu) \in \mathbb{D}_0$, one has*

$$U_{s,k}^B(\mathfrak{a}) \ll p^{B\varepsilon} U_{s,k}^{B,H}(\mathfrak{a}).$$

The proof of Theorem 15.4 and its corollary follow just as in the corresponding argument of §§3-10 above in the rational case. Given our adjustments in notation, the argument follows verbatim, provided that the reader exercises care in ensuring that these notational perturbations are correctly administered. We therefore move on directly to consider the proof of Theorem 15.1.

The proof of Theorem 15.1. In all essentials, the proof of Theorem 15.1 follows the proof of Theorem 1.1 given in §11 and §12, mutatis mutandis, and so we shall be very brief concerning the details. Here, with the notational modifications in hand, the only parts of these sections that require further discussion are located in §12. Let $\mathcal{Z}$ denote the set of zeros of $W(t; \boldsymbol{\varphi})$ lying in $\mathfrak{O}_K$, and let X be sufficiently large in terms of $\boldsymbol{\varphi}$, k, ε, K and Ω. We suppose that $s = k(k+1)/2$. We note that one again has

$$\mathrm{card}(\mathcal{Z}) \leqslant \deg(W(t; \boldsymbol{\varphi})) \ll 1.$$

Define

$$F(\boldsymbol{\alpha}; X) = \rho_0^{-1} \sum_{\nu \in \mathcal{B}(X)} \mathfrak{a}_\nu E(\psi(\nu; \boldsymbol{\alpha}))$$

and

$$F_0(\boldsymbol{\alpha}; X) = \rho_0^{-1} \sum_{\substack{\nu \in \mathcal{B}(X) \\ \nu \notin \mathcal{Z}}} \mathfrak{a}_\nu E(\psi(\nu; \boldsymbol{\alpha})).$$

We may suppose that the sequence $(\mathfrak{a}_\nu)$ satisfies the property that $\mathfrak{a}_\nu = 0$ whenever $\nu \notin \mathcal{B}(X)$. Our task is to cover the exponential sum $F_0(\boldsymbol{\alpha}; X)$ by analogous exponential sums with variables constrained by appropriate non-singularity conditions modulo $\mathfrak{p}$, for suitable prime ideals $\mathfrak{p} = (\pi)$ in $\mathfrak{O}_K$.

Given a solution $\mathbf{x}, \mathbf{y} \in (\mathcal{B}(X) \setminus \mathcal{Z})^s$ of the system (1.4), the algebraic integer

$$\prod_{i=1}^{s} W(x_i; \boldsymbol{\varphi}) W(y_i; \boldsymbol{\varphi})$$

is non-zero, and has norm bounded above by CX^D, for some $C > 0$ depending at most on K, s and the coefficients of $\boldsymbol{\varphi}$, and D a positive integer with

$$D \leqslant 2s \sum_{j=1}^{k} \deg(\varphi_j).$$

Let $\mathcal{P}$ denote the set of elements $\pi \in \mathfrak{O}_K$ having the property that (π) is a prime ideal satisfying the condition

$$(\log X)^2 < \mathrm{N}(\pi) \leqslant 3(\log X)^2.$$

By the prime ideal theorem in number fields (see [28]), when X is sufficiently large, the number of such elements is at least $\frac{1}{2}(\log X)^2/\log\log X$. One therefore has

$$\mathrm{N}\Big(\prod_{\pi\in\mathcal{P}} \pi\Big) > (\log X)^{(\log X)^2/\log\log X} > (CX^D)^n.$$

Thus we deduce that for each solution $\mathbf{x}, \mathbf{y} \in (\mathcal{B}(X) \setminus \mathcal{Z})^s$ of (1.4) counted by the mean value

$$\oint |F_0(\boldsymbol{\alpha}; X)|^{2s} \,\mathrm{d}\boldsymbol{\alpha},$$

there exists $\pi \in \mathcal{P}$ with

$$\prod_{i=1}^{s} W(x_i; \boldsymbol{\varphi}) W(y_i; \boldsymbol{\varphi}) \not\equiv 0 \pmod{\mathfrak{p}},$$

in which we write $\mathfrak{p} = (\pi)$. In particular, one has

$$\oint |F_0(\boldsymbol{\alpha}; X)|^{2s} \,\mathrm{d}\boldsymbol{\alpha} \leqslant \sum_{\pi\in\mathcal{P}} \oint |F_\pi(\boldsymbol{\alpha}; X)|^{2s} \,\mathrm{d}\boldsymbol{\alpha}, \tag{15.3}$$

where

$$F_\pi(\boldsymbol{\alpha}; X) = \rho_0^{-1} \sum_{\substack{\nu\in\mathcal{B}(X)\\ W(\nu;\boldsymbol{\varphi})\not\equiv 0 \pmod{\mathfrak{p}}}} |\mathfrak{a}_\nu| \, E(\psi(\nu; \boldsymbol{\alpha})).$$

The argument of §12 now resumes. Let $\tau > 0$ be sufficiently small in terms of s and k. We take

$$B = \left\lceil \frac{k \log(pX)}{\log p} \right\rceil, \quad c = \lceil \tau B \rceil \quad \text{and} \quad H = \lceil B/k \rceil - c,$$

which ensures that $pX \leqslant \mathrm{N}(\pi^{H+c}) \leqslant p^2 X$. Then we may suppose that B is sufficiently large in terms of τ, as well as s, k and ε. By orthogonality, one has

$$\oint |F_\pi(\boldsymbol{\alpha}; X)|^{2s} \,\mathrm{d}\boldsymbol{\alpha} \leqslant \oint_{\mathfrak{p}^B} |F_\pi(\boldsymbol{\alpha}; X)|^{2s} \,\mathrm{d}\boldsymbol{\alpha}.$$

Thus, as a consequence of Corollary 15.5, just as in the argument of the proof of Theorem 11.1 leading to (11.9), one finds that there is a $\mathfrak{p}^c$-spaced system $\boldsymbol{\Psi}$ and complex sequence $\mathfrak{c}$ with $|\mathfrak{c}_y| = |\mathfrak{a}_{\pi^c y+\xi}|$ for which

$$\oint_{\mathfrak{p}^B} |F_\pi(\boldsymbol{\alpha}; X)|^{2s} \,\mathrm{d}\boldsymbol{\alpha} \ll p^{sc+B\varepsilon} \rho_0(\mathfrak{a})^{-2} \sum_{\xi \bmod \mathfrak{p}^c} \rho_c(\xi)^2 U_{s,k}^{B-kc,H,\boldsymbol{\Psi}}(\mathfrak{c}).$$

Notice that whenever $\nu, \nu' \in \mathcal{B}(X)$, one has $\nu - \nu' \in \mathcal{B}(2X)$, and hence $\mathrm{N}(\nu - \nu') \ll X$. Then since X is sufficiently large and $p > (\log X)^2$, we obtain the implication

$$\nu \equiv \nu' \pmod{\mathfrak{p}^{H+c}} \quad \Rightarrow \quad \nu = \nu'. \tag{15.4}$$

Consequently, just as in the concluding paragraph of the proof of Theorem 11.1, we find that

$$U_{s,k}^{B-kc,H,\mathbf{\Psi}}(\mathfrak{c}) \ll (1 + X/p^{c+H})^s \ll 1.$$

Thus we deduce that

$$\oint_{\mathfrak{p}^B} |F_\pi(\boldsymbol{\alpha}; X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \ll p^{sc+B\varepsilon} \rho_0(\mathfrak{a})^{-2} \sum_{\xi \bmod \mathfrak{p}^c} \rho_c(\xi)^2 \ll p^{(2s\tau+\varepsilon)B}.$$

Since τ was chosen sufficiently small in terms of s and k, it follows that for each positive number δ, one has

$$\oint |F_\pi(\boldsymbol{\alpha}; X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \ll X^\delta,$$

and thus, on recalling (15.3),

$$\begin{aligned} \oint |F(\boldsymbol{\alpha}; X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} &\ll 1 + \oint |F_0(\boldsymbol{\alpha}; X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \\ &\ll 1 + X^\delta \sum_{\pi \in \mathcal{P}} 1 \ll X^{2\delta}. \end{aligned}$$

The conclusion of the theorem follows on recalling the definitions of $F(\boldsymbol{\alpha}; X)$ and the weight ρ_0. □

It is difficult to resist announcing an easy consequence of Theorem 15.1 that follows by a straightforward application of the circle method in number fields. This supplies a Hasse principle for diagonal forms.

Theorem 15.6. *Let K be an algebraic extension of $\mathbb{Q}$ of finite degree. Let $k, s \in \mathbb{N}$, and suppose that $s \geqslant k^2 + k + 1$. Suppose also that $a_1, \dots, a_s \in K$, and that the equation*

$$a_1 x_1^k + \dots + a_s x_s^k = 0$$

has non-zero solutions in every completion K_v of K. Then this equation has a solution $\mathbf{x} \in K^s \setminus \{\mathbf{0}\}$.

Proof. Write

$$G(\boldsymbol{\alpha}; X) = \sum_{\nu \in \mathcal{B}(X)} E(\alpha_1 \nu + \dots + \alpha_k \nu^k).$$

Then it follows from the triangle inequality that

$$\begin{aligned} \int_{\mathbb{T}} \Big| \sum_{\nu \in \mathcal{B}(X)} E(\alpha \nu^k) \Big|^{2s} \, \mathrm{d}\alpha &= \sum_{\mathbf{h}} \int_{\mathbb{T}^k} |G(\boldsymbol{\alpha}; X)|^{2s} E(-\alpha_1 h_1 - \dots - \alpha_{k-1} h_{k-1}) \, \mathrm{d}\boldsymbol{\alpha} \\ &\ll X^{k(k-1)/2} J_{s,k}(X; K), \end{aligned}$$

in which the $(k-1)$-tuples $\mathbf{h}$ are summed over the boxes

$$h_j \in 2sC\mathcal{B}(X^j) \quad (1 \leqslant j \leqslant k-1),$$

for a positive number C sufficiently large in terms of Ω and k. Thus, when $s \geqslant k(k+1)$, one finds from Corollary 15.3 that

$$\int_{\mathbb{T}} \Big| \sum_{\nu \in \mathcal{B}(X)} E(\alpha \nu^k) \Big|^{2s} \mathrm{d}\alpha \ll X^{2s-k+\varepsilon}.$$

Using this estimate as a substitute for [4, Lemma 2] in the proof of [4, Theorem 3], the proof of the theorem follows via a standard application of the circle method in algebraic number fields. □

Corollary 15.7. *Let L be an algebraic extension of $\mathbb{Q}$, possibly of infinite degree. Let $k, s \in \mathbb{N}$ with k odd, and suppose that $s \geqslant \exp(8(\log k)^2)$. Suppose also that $a_1, \dots, a_s \in L$. Then the equation*

$$a_1 x_1^k + \dots + a_s x_s^k = 0 \tag{15.5}$$

has a solution $\mathbf{x} \in L^s \setminus \{\mathbf{0}\}$. When L is a totally imaginary extension of $\mathbb{Q}$, the same conclusion holds also for even k.

Proof. Since L is an algebraic extension of $\mathbb{Q}$, the coefficients $a_1, \dots, a_s$ are algebraic. Put $K = \mathbb{Q}(a_1, \dots, a_s)$. Then $K : \mathbb{Q}$ is a finite algebraic extension of $\mathbb{Q}$, and it follows from [9, Theorem 1], just as in the proof of [57, Theorem 4.4], that in every completion K_v of K the equation (15.5) has a non-zero solution. Note that when the place v is infinite, this is trivially inferred from the hypothesis that, either k is odd, or else L is a totally imaginary extension of $\mathbb{Q}$ and k is even. It therefore follows from Theorem 15.6 that the equation (15.5) possesses a non-zero K-rational solution, and hence also a non-zero L-rational solution. □

Corollary 15.8. *Let L be a totally imaginary algebraic extension of $\mathbb{Q}$, possibly of infinite degree. Let $k, s, r \in \mathbb{N}$ with $k \geqslant 3$, and suppose that*

$$s > r^{2^{k-1}} \exp(2^{k+2}(\log k)^2).$$

Then, whenever $F_i \in L[x_1, \dots, x_s]$ $(1 \leqslant i \leqslant r)$ are homogeneous of degree k, the system of equations

$$F_i(\mathbf{x}) = 0 \quad (1 \leqslant i \leqslant r)$$

possess a simultaneous solution $\mathbf{x} \in L^s \setminus \{\mathbf{0}\}$.

Proof. One may follow the argument of the proof of [46, Corollary 1.3], mutatis mutandis, noting the proof of [57, Theorem 4.4], and substituting Corollary 15.7 into [46, Theorem 1]. □

The conclusion of this corollary provides an explicit version of a theorem of Peck [31]. We note that [46, Corollary 1.3] establishes a similar conclusion subject to the stronger constraint $s > r^{2^{k-1}} \exp(2^k k)$. We intend to explore further applications of Theorem 15.1 in a later paper.

16. Multidimensional analogues via restriction of scalars

The conclusions of §15 may be employed to establish a class of mean value estimates associated with multidimensional systems. In simplest terms, what we have in mind is that a mean value estimate of the shape given in Corollary 15.2, may be reinterpreted as a multidimensional mean value estimate over a lower degree field extension of $\mathbb{Q}$. Perhaps this is best illustrated with a concrete example. Thus, consider the Vinogradov system of degree 3 over $K = \mathbb{Q}(\sqrt{-2})$. Theorem 15.1 shows that whenever $s \geqslant 1$, the number $J_{s,3}(X;K)$ of solutions of the system

$$\sum_{i=1}^{s}(x_i^j - y_i^j) = 0 \quad (1 \leqslant j \leqslant 3), \tag{16.1}$$

with $\mathbf{x}, \mathbf{y} \in \mathcal{B}(X)^s$, satisfies

$$J_{s,3}(X;K) \ll X^{\varepsilon}(X^s + X^{2s-6}).$$

However, for each such solution $\mathbf{x}, \mathbf{y}$, one may write

$$x_i = u_i + v_i\sqrt{-2} \quad \text{and} \quad y_i = z_i + w_i\sqrt{-2},$$

with $\mathbf{u}, \mathbf{v}, \mathbf{z}, \mathbf{w} \in [-\frac{1}{2}X^{1/2}, \frac{1}{2}X^{1/2})^s \cap \mathbb{Z}^s$. Define

$$\phi_3(x,y) = x^3 - 6xy^2, \quad \psi_3(x,y) = 3x^2y - 2y^3,$$
$$\phi_2(x,y) = x^2 - 2y^2, \quad \psi_2(x,y) = xy,$$
$$\phi_1(x,y) = x, \quad \psi_1(x,y) = y.$$

Then by expanding the expressions

$$(u_i + v_i\sqrt{-2})^j \quad \text{and} \quad (z_i + w_i\sqrt{-2})^j,$$

for $1 \leqslant j \leqslant 3$, and writing the result in terms of the integral coordinate basis $\{1, \sqrt{-2}\}$, one sees that (16.1) holds if and only if the system

$$\sum_{i=1}^{s}(\phi_j(u_i,v_i) - \phi_j(z_i,w_i)) = 0,$$
$$\sum_{i=1}^{s}(\psi_j(u_i,v_i) - \psi_j(z_i,w_i)) = 0,$$

is satisfied simultaneously for $1 \leqslant j \leqslant 3$.

Denote by $J_s(Y;\boldsymbol{\varphi},\boldsymbol{\psi})$ the number of integral solutions of the latter system with $1 \leqslant \mathbf{u}, \mathbf{v}, \mathbf{z}, \mathbf{w} \leqslant Y$. Then from the estimate

$$J_{s,3}(X;\mathbb{Q}(\sqrt{-2})) \ll X^{\varepsilon}(X^s + X^{2s-6}),$$

available via Corollary 15.3, we deduce that

$$J_s(Y;\boldsymbol{\varphi},\boldsymbol{\psi}) \ll Y^{\varepsilon}(Y^{2s} + Y^{4s-12}).$$

This estimate delivers the main conjecture for this two dimensional system.

In order to describe this phenomenon in wider generality, we introduce some notation. Let K be an algebraic extension of $\mathbb{Q}$ with $[K : \mathbb{Q}] = d$.

Let L be an algebraic extension of K with $[L : K] = n$, and let $\mathfrak{O}_{L/K}$ denote the ring of integers associated with the field extension $L : K$. Write $\{\omega_1, \dots, \omega_n\}$ for an $\mathfrak{O}_K$-integral coordinate basis of $L : K$. Consider polynomials $\varphi_1, \dots, \varphi_k \in \mathfrak{O}_{L/K}[t]$ with $W(t; \boldsymbol{\varphi}) \neq 0$. We say that the system of polynomials $\psi_{lj}(t_1, \dots, t_n) \in \mathfrak{O}_{K/\mathbb{Q}}[\mathbf{t}]$ $(1 \leqslant l \leqslant n)$ is *generated from* $\boldsymbol{\varphi}$ *by restriction of* L *down to* K when, for some $\lambda_{i1}, \dots, \lambda_{in} \in K$, with $\det(\lambda_{il})_{1 \leqslant i,l \leqslant n} \neq 0$, one has

$$\varphi_j(t_1\omega_1 + \dots + t_n\omega_n) = \sum_{i=1}^{n} \omega_i \sum_{l=1}^{n} \lambda_{il}\psi_{lj}(t_1, \dots, t_n) \quad (1 \leqslant j \leqslant k).$$

We are now equipped to announce an analogue of the second conclusion of Theorem 1.1 for certain multidimensional systems. Here, we interpret $\mathcal{B}(X) = \mathcal{B}_K(X)$, as before, as a subset of $\mathfrak{O}_{K/\mathbb{Q}}$ relative to a fixed coordinate basis for $\mathfrak{O}_{K/\mathbb{Q}}$ over $\mathbb{Z}$.

Theorem 16.1. *Let $L : K : \mathbb{Q}$ be a tower of algebraic field extensions with $[L : \mathbb{Q}] < \infty$ and $[L : K] = n$. Given a system of polynomials $\varphi_1, \dots, \varphi_k \in \mathfrak{O}_{L/K}[t]$ with $W(t; \boldsymbol{\varphi}) \neq 0$, suppose that*

$$\psi_{lj}(t_1, \dots, t_n) \in \mathfrak{O}_{K/\mathbb{Q}}[\mathbf{t}] \quad (1 \leqslant l \leqslant n,\ 1 \leqslant j \leqslant k)$$

is generated from $\boldsymbol{\varphi}$ by restriction of L down to K. Finally, suppose that $1 \leqslant s \leqslant k(k+1)/2$ and $\varepsilon > 0$. Then the number $J_s(X; \boldsymbol{\psi}; K)$ of solutions of the system

$$\sum_{i=1}^{s} (\psi_{lj}(x_{i1}, \dots, x_{in}) - \psi_{lj}(y_{i1}, \dots, y_{in})) = 0 \quad (1 \leqslant l \leqslant n, 1 \leqslant j \leqslant k), \quad (16.2)$$

with $\mathbf{x}, \mathbf{y} \in \mathcal{B}(X)^{ns}$, satisfies $J_s(X; \boldsymbol{\psi}; K) \ll (\mathrm{card}(\mathcal{B}(X)))^{ns+\varepsilon}$.

Proof. The system of equations (16.2) over $\mathfrak{O}_{K/\mathbb{Q}}$ is satisfied if and only if the system

$$\sum_{i=1}^{s} (\varphi_j(u_i) - \varphi_j(v_i)) = 0 \quad (1 \leqslant j \leqslant k), \quad (16.3)$$

is satisfied over $\mathfrak{O}_{L/K}$ by

$$u_i = x_{i1}\omega_1 + \dots + x_{in}\omega_n \quad \text{and} \quad v_i = y_{i1}\omega_1 + \dots + y_{in}\omega_n \quad (1 \leqslant i \leqslant s).$$

By Corollary 15.2, when $1 \leqslant s \leqslant k(k+1)/2$ and $\varepsilon > 0$, the total number T of solutions of (16.3) with

$$\mathbf{u}, \mathbf{v} \in \{r_1\omega_1 + \dots + r_n\omega_n : \mathbf{r} \in \mathcal{B}(X)^n\}^s$$

satisfies

$$T \ll X^{ns+\varepsilon} \ll (\mathrm{card}(\mathcal{B}(X)))^{ns+\varepsilon}.$$

The conclusion of the theorem follows. □

This theorem supplies infinitely many examples of multidimensional systems for which the main conjecture holds. We have merely to examine systems of polynomials generated from $(t, t^2, \dots, t^k)$ by restriction of one number field to a subfield (perhaps $\mathbb{Q}$). Other multidimensional conclusions can be found

in [29], where the authors aimed for completely general (though non-optimal) results for arbitrary translation-dilation invariant systems in many variables. The most recent work based on decoupling in certain two dimensional problems may be found in [7, 17].

17. Vinogradov's mean value theorem in function fields

One of the key messages to be extracted from this memoir is that the nested efficient congruencing method is sufficiently robust that it may be employed in a myriad environments with minimal adjustment. In particular, in contrast with the l^2-decoupling method of [8], we require no multilinear Kakeya estimates that might be unavailable, or even inherently mysterious in nature, when contemplated in different settings. Nested efficient congruencing is a method likely to achieve success for analogues of Vinogradov's mean value theorem involving discrete sets of points in any Henselian field. All that one requires are appropriate characters with which to engineer certain orthogonality relations. One need not be constrained to non-archimedean environments, moreover, since with additional effort involved in handling non-ultrametric inequalities, the same arguments apply equally well in archimedean environments such as the real or complex numbers. Indeed, the use of real short intervals underpinned the original approach of Vinogradov [40] (see also [22], [37, Chapter VI], and [36] for recent developments associated with efficient congruencing). In this section we illustrate this robustness with one final example, namely that of function fields.

We explain the consequences of the work of §§3-12 in the most basic situation of a function field $\mathbb{F}_q(t)$ of characteristic $p > k$, where k is the number of equations at hand. We emphasise that this is very far from the strongest type of result available in the function field setting, since the situation with small characteristic $p \leqslant k$ is considerably more complicated. Comprehensive conclusions achieving the main conjecture in Vinogradov's mean value theorem are the subject of work in progress by the author joint with Y.-R. Liu. In this section we avoid intruding on the latter work, and instead extract only the results that may be obtained with essentially no effort from our analysis in §§3-12.

We begin by recalling the infrastructure required for harmonic analysis in the function field setting. Our coefficients come from the finite field $\mathbb{F}_q$ of characteristic p having $q = p^l$ elements. Associated with the polynomial ring $\mathbb{O} = \mathbb{F}_q[t]$ defined over the field $\mathbb{F}_q$ is its field of fractions $\mathbb{K} = \mathbb{F}_q(t)$. In this section, we take d to be the main parameter, a sufficiently large natural number, and we put

$$\mathbb{O}_d = \{\nu \in \mathbb{F}_q[t] : \deg(\nu) \leqslant d\}.$$

We write $\mathbb{K}_\infty = \mathbb{F}_q((1/t))$ for the completion of $\mathbb{F}_q(t)$ at ∞. One may write each element $\alpha \in \mathbb{K}_\infty$ in the shape $\alpha = \sum_{i \leqslant n} a_i t^i$ for some $n \in \mathbb{Z}$ and coefficients $a_i = a_i(\alpha)$ in $\mathbb{F}_q$ $(i \leqslant n)$. We define $\operatorname{ord} \alpha$ to be the largest integer i for which $a_i(\alpha) \neq 0$. We then write $\langle \alpha \rangle$ for $q^{\operatorname{ord} \alpha}$. In this context, we adopt the

convention that $\operatorname{ord} 0 = -\infty$ and $\langle 0\rangle = 0$. Consider next the compact additive subgroup $\mathbb{T}$ of $\mathbb{K}_\infty$ defined by $\mathbb{T} = \{\alpha \in \mathbb{K}_\infty : \langle\alpha\rangle < 1\}$. Every element α of $\mathbb{K}_\infty$ can be written uniquely in the shape $\alpha = [\alpha] + \|\alpha\|$, where $[\alpha] \in \mathbb{F}_q[t]$ and $\|\alpha\| \in \mathbb{T}$, and we may normalise any Haar measure $\mathrm{d}\alpha$ on $\mathbb{K}_\infty$ in such a manner that $\int_{\mathbb{T}} 1\,\mathrm{d}\alpha = 1$.

We are now equipped to define an analogue of the exponential function. There is a non-trivial additive character $e_q : \mathbb{F}_q \to \mathbb{C}^\times$ defined for each $a \in \mathbb{F}_q$ by taking $e_q(a) = e(\mathrm{tr}(a)/p)$, where we write $e(z)$ for $e^{2\pi iz}$, and where $\mathrm{tr}:\mathbb{F}_q \to \mathbb{F}_p$ denotes the familiar trace map. This character induces a map $e : \mathbb{K}_\infty \to \mathbb{C}^\times$ by defining, for each element $\alpha \in \mathbb{K}_\infty$, the value of $e(\alpha)$ to be $e_q(a_{-1}(\alpha))$. The orthogonality relation underlying the Fourier analysis of $\mathbb{F}_q[t]$, established for example in [26, Lemma 1], takes the shape

$$\int_{\mathbb{T}} e(h\alpha)\,d\alpha = \begin{cases} 0, & \text{when } h \in \mathbb{F}_q[t] \setminus \{0\}, \\ 1, & \text{when } h = 0. \end{cases} \tag{17.1}$$

Theorem 17.1. *Suppose that $\varphi_j \in \mathbb{O}[x]$ $(1 \leqslant j \leqslant k)$ is a system of polynomials with $W(x;\boldsymbol{\varphi}) \neq 0$. Let s be a positive real number with $s \leqslant k(k+1)/2$. Also, suppose that $(\mathfrak{a}_\nu)_{\nu\in\mathbb{O}}$ is a sequence of complex numbers. Then provided that $\mathrm{ch}(\mathbb{F}_q) > k$ and $\varepsilon > 0$, one has*

$$\int_{\mathbb{T}^k}\Big|\sum_{\nu\in\mathbb{O}_d}\mathfrak{a}_\nu e(\alpha_1\varphi_1(\nu)+\ldots+\alpha_k\varphi_k(\nu))\Big|^{2s}\,\mathrm{d}\boldsymbol{\alpha} \ll (q^d)^\varepsilon\Big(\sum_{\nu\in\mathbb{O}_d}|\mathfrak{a}_\nu|^2\Big)^s.$$

In particular, one has

$$\int_{\mathbb{T}^k}\Big|\sum_{\nu\in\mathbb{O}_d} e(\alpha_1\varphi_1(\nu)+\ldots+\alpha_k\varphi_k(\nu))\Big|^{2s}\,\mathrm{d}\boldsymbol{\alpha} \ll (q^d)^{s+\varepsilon}.$$

In this section, implicit constants in Vinogradov's notation may depend on s, k, $\boldsymbol{\varphi}$, q, and also the small positive number ε. As an immediate consequence of the orthogonality relation (17.1), one obtains the following corollary.

Corollary 17.2. *When $s \in \mathbb{N}$, denote by $N_{s,\boldsymbol{\varphi}}(d,q)$ the number of solutions of the system of equations*

$$\sum_{i=1}^{s}(\varphi_j(x_i) - \varphi_j(y_i)) = 0 \quad (1 \leqslant j \leqslant k),$$

with $x_i, y_i \in \mathbb{O}_d$ $(1 \leqslant i \leqslant s)$. Suppose that $\mathrm{ch}(\mathbb{F}_q) > k$. Then whenever $W(x;\boldsymbol{\varphi}) \neq 0$ and $s \leqslant k(k+1)/2$, one has $N_{s,\boldsymbol{\varphi}}(d,q) \ll (q^d)^{s+\varepsilon}$.

Finally, when $k \in \mathbb{N}$ and $s > 0$, we define

$$J_{s,k}(d;\mathbb{F}_q) = \int_{\mathbb{T}^k}\Big|\sum_{\nu\in\mathbb{O}_d} e(\alpha_1\nu+\ldots+\alpha_k\nu^k)\Big|^{2s}\,\mathrm{d}\boldsymbol{\alpha}.$$

Theorem 17.1 delivers an analogue of the main conjecture in Vinogradov's mean value theorem in function fields of large characteristic.

Corollary 17.3. *Suppose that $k \in \mathbb{N}$ and $s > 0$. Then whenever* $\mathrm{ch}(\mathbb{F}_q) > k$ *and $\varepsilon > 0$, one has*

$$J_{s,k}(d;\mathbb{F}_q) \ll (q^d)^{\varepsilon}\left((q^d)^s + (q^d)^{2s-k(k+1)/2}\right).$$

There is an unpublished manuscript, more than a decade old, of Y.-R. Liu and the author in the function field setting that has been quoted from time to time, and was updated to reflect developments arising from the early efficient congruencing methods. This was subsumed by the multidimensional work [27], which would achieve an analogue of Corollary 17.3 for $s > k(k+1)$ when $\mathrm{ch}(\mathbb{F}_q) > k$, with sharper conclusions available for $\mathrm{ch}(\mathbb{F}_q) \leqslant k$. Forthcoming work of Y.-R. Liu and the author removes the hypothesis $\mathrm{ch}(\mathbb{F}_q) > k$ from an analogue of Corollary 17.3 in which the main conjecture is proved in general for Vinogradov's mean value theorem in function fields. This forthcoming joint work represents the definitive statement on the subject.

The conclusion of Theorem 17.1 follows from an analogue of Theorem 3.1, as we now sketch. The details are again strikingly similar to those in the situation over $\mathbb{Z}$ described in §§3–12, so we are skimpy on details. Once more, the parameter k is an integer with $k \geqslant 1$, and we consider polynomials $\varphi_j \in \mathbb{O}[x]$ $(1 \leqslant j \leqslant k)$. Throughout the argument, the ring of polynomials $\mathbb{O}$ replaces $\mathbb{Z}$. Let $\pi \in \mathbb{O}$ be a monic irreducible polynomial, and hence of positive degree. The definitions of §3 must be made once again, mutatis mutandis, where we replace $\mathbb{Z}$ by $\mathbb{O}$, congruences modulo p^h by congruences modulo π^h, and the function $e(z)$ from §3 by its doppelgänger defined in this section, throughout. The definitions (1.9) and (3.5) must again be adjusted to the present setting. First, when $F : \mathbb{T}^k \to \mathbb{C}$ is integrable, we write

$$\oint F(\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha} = \int_{\mathbb{T}^k} F(\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha}.$$

Next, when B is a positive integer, we define

$$\oint_{\pi^B} F(\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha} = \langle \pi\rangle^{-kB} \sum_{u_1 \bmod \pi^B} \dots \sum_{u_k \bmod \pi^B} F(\mathbf{u}\pi^{-B}),$$

where the summations are taken over complete sets of residues modulo π^B. With this definition, one may verify that for $\nu \in \mathbb{O}$, one has the orthogonality relation

$$\oint_{\pi^B} e(\alpha\nu)\,\mathrm{d}\alpha = \begin{cases} 1, & \text{when } \nu \equiv 0 \pmod{\pi^B},\\ 0, & \text{when } \nu \not\equiv 0 \pmod{\pi^B}.\end{cases}$$

These definitions permit an analogue of Theorem 3.1 to be stated in the function field setting. We recall the definitions (3.6) and (3.8), now adjusted to their function field manifestations.

Theorem 17.4. *Let $\mathbb{O} = \mathbb{F}_q[t]$. Suppose that $k \in \mathbb{N}$, and that $\pi \in \mathbb{O}$ is a monic irreducible polynomial. Then, under the assumption that* $\mathrm{ch}(\mathbb{F}_q) > k$, *one has $\lambda(k(k+1)/2, k) = 0$.*

Corollary 17.5. *Let $\mathbb{O} = \mathbb{F}_q[t]$. Suppose that $k \in \mathbb{N}$, and that $\pi \in \mathbb{O}$ is a monic irreducible polynomial. Suppose also that $\tau > 0$ and $\varepsilon > 0$. Let B be sufficiently large in terms of k, τ and ε. Put $s = k(k+1)/2$ and $H = \lceil B/k \rceil$. Finally, suppose that $\mathrm{ch}(\mathbb{F}_q) > k$. Then for every $\boldsymbol{\varphi} \in \Phi_\tau(B)$ and every sequence $(\mathfrak{a}_n) \in \mathbb{D}_0$, one has*

$$U_{s,k}^B(\mathfrak{a}) \ll \langle \pi \rangle^{B\varepsilon} U_{s,k}^{B,H}(\mathfrak{a}).$$

The proof of Theorem 17.4 and its corollary follow just as in the corresponding argument of §§3–10 above in the rational integer case. Given our adjustments in notation, the argument follows verbatim, provided that the reader exercises due diligence in ensuring that these notational perturbations are correctly construed. Perhaps it is worth noting that the restriction to situations subject to the condition $\mathrm{ch}(\mathbb{F}_q) > k$ arises from the implied assumption that $W(x;\boldsymbol{\varphi}) \neq 0$. If the system $\boldsymbol{\varphi}$ is π^c-spaced for some positive integer c and monic irreducible polynomial π, then the determinant $W(x;\boldsymbol{\varphi})$ is congruent modulo π^c to a triangular determinant with entries $j!$ $(1 \leqslant j \leqslant k)$ along the diagonal. When $\mathrm{ch}(\mathbb{F}_q) \leqslant k$, it follows that one of these diagonal entries is zero modulo π, and hence one fails to be able to engineer a non-vanishing Wronskian modulo π. The solution to this problem is to make use of a basis for Taylor expansions in small positive characteristic smaller than the naïve basis $\{1, x, x^2, \dots\}$ of consecutive powers. This then entails adjusting also the definition of the Wronskian for function fields accordingly, and in this setting the environment has changed sufficiently that other adjustments are required in the discussion of §§3-10. This is a matter to which we return in our forthcoming joint work with Y.-R. Liu.

Having left the reader with the mechanical reproduction of this verbatim proof, we move on immediately to the proof of Theorem 17.1

The proof of Theorem 17.1. Our argument follows the proof of Theorem 1.1 given in §§11 and 12 with very few adjustments. We again denote by $\mathcal{Z}$ the set of zeros of $W(x;\boldsymbol{\varphi})$ lying in $\mathbb{O}$. Since $W(x;\boldsymbol{\varphi})$ is non-zero, it follows that

$$\mathrm{card}(\mathcal{Z}) \leqslant \deg(W(x;\boldsymbol{\varphi})) \ll 1.$$

We define

$$F(\boldsymbol{\alpha}; d) = \rho_0^{-1} \sum_{\nu \in \mathbb{O}_d} \mathfrak{a}_\nu e(\psi(\nu;\boldsymbol{\alpha}))$$

and

$$F_0(\boldsymbol{\alpha}; d) = \rho_0^{-1} \sum_{\nu \in \mathbb{O}_d \setminus \mathcal{Z}} \mathfrak{a}_\nu e(\psi(\nu;\boldsymbol{\alpha})).$$

We may suppose that the sequence $(\mathfrak{a}_\nu)$ satisfies the property that $\mathfrak{a}_\nu = 0$ whenever $\nu \notin \mathbb{O}_d$. Our task is to cover the exponential sum $F_0(\boldsymbol{\alpha}; d)$ by exponential sums with variables constrained by non-singularity conditions modulo π, for suitable monic irreducible polynomials $\pi \in \mathbb{O}$.

Given a solution $\mathbf{x}, \mathbf{y} \in (\mathbb{O}_d \setminus \mathcal{Z})^s$ of the system (1.4), the polynomial

$$\prod_{i=1}^{s} W(x_i; \boldsymbol{\varphi}) W(y_i; \boldsymbol{\varphi})$$

is non-zero, and has degree bounded above by $C + Dd$, for some $C > 0$ depending at most on s and the coefficients of $\boldsymbol{\varphi}$, and D a positive integer with

$$D \leqslant 2s \sum_{j=1}^{k} \deg(\varphi_j).$$

Let $\mathcal{P}$ denote the set of elements $\pi \in \mathbb{O}$ with π monic and irreducible, and satisfying

$$(\log d)^2 < \deg(\pi) \leqslant 3(\log d)^2.$$

By the analogue of the prime number theorem over $\mathbb{F}_q[t]$, when d is sufficiently large, the number of such elements is at least $q^{2(\log d)^2}/(\log d)^2$. One therefore has

$$\deg\Big(\prod_{\pi \in \mathcal{P}} \pi \Big) = \sum_{\pi \in \mathcal{P}} \deg(\pi) > q^{2(\log d)^2} > d^{\log d}.$$

Thus $\deg\big(\prod_{\pi \in \mathcal{P}} \pi\big) > C + Dd$. Then for each solution $\mathbf{x}, \mathbf{y} \in \mathbb{O}^s$ of (1.4) counted by the mean value

$$\oint |F_0(\boldsymbol{\alpha}; d)|^{2s} \, \mathrm{d}\boldsymbol{\alpha},$$

there exists $\pi \in \mathcal{P}$ with

$$\prod_{i=1}^{s} W(x_i; \boldsymbol{\varphi}) W(y_i; \boldsymbol{\varphi}) \not\equiv 0 \pmod{\pi}.$$

In particular, one has

$$\oint |F_0(\boldsymbol{\alpha}; d)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \leqslant \sum_{\pi \in \mathcal{P}} \oint |F_\pi(\boldsymbol{\alpha}; d)|^{2s} \, \mathrm{d}\boldsymbol{\alpha}, \tag{17.2}$$

where

$$F_\pi(\boldsymbol{\alpha}; d) = \rho_0^{-1} \sum_{\nu \in \mathbb{O}_d \setminus \mathcal{Z}} |\mathfrak{a}_\nu| e(\psi(\nu; \boldsymbol{\alpha})).$$

Resuming the argument of §12, we take $\tau > 0$ to be sufficiently small in terms of s and k, and put

$$B = \lceil k(d+1)/\deg(\pi) \rceil, \quad c = \lceil \tau B \rceil \quad \text{and} \quad H = \lceil B/k \rceil - c,$$

which ensures that

$$d + 1 \leqslant \deg\big(\pi^{H+c}\big) \leqslant d + 1 + \deg(\pi).$$

Then we may suppose that B is sufficiently large in terms of τ, as well as s, k and ε. By orthogonality, one has

$$\oint |F_\pi(\boldsymbol{\alpha}; d)|^{2s} \, \mathrm{d}\boldsymbol{\alpha} \leqslant \oint_{\pi^B} |F_\pi(\boldsymbol{\alpha}; d)|^{2s} \, \mathrm{d}\boldsymbol{\alpha}.$$

Then, by Corollary 17.5, just as in the argument of the proof of Theorem 11.1 leading to (11.9), one sees that there is a π^c-spaced system $\boldsymbol{\Psi}$ and complex sequence $\mathfrak{c}$ with $|\mathfrak{c}_y| = |\mathfrak{a}_{\pi^c y+\xi}|$ for which one has

$$\oint_{\pi^B} |F_\pi(\boldsymbol{\alpha}; d)|^{2s}\, \mathrm{d}\boldsymbol{\alpha} \ll \langle \pi \rangle^{sc+B\varepsilon} \rho_0(\mathfrak{a})^{-2} \sum_{\xi \bmod \pi^c} \rho_c(\xi)^2 U_{s,k}^{B-kc,H,\boldsymbol{\Psi}}(\mathfrak{c}).$$

Observe that when $\nu, \nu' \in \mathbb{O}_d$, then

$$\nu \equiv \nu' \pmod{\pi^{H+c}} \quad \Rightarrow \quad \nu = \nu'.$$

Thus, as in the concluding paragraph of the proof of Theorem 11.1, one infers that

$$U_{s,k}^{B-kc,H,\boldsymbol{\Psi}}(\mathfrak{c}) \ll (1 + q^d/\langle \pi \rangle^{H+c})^s \ll 1.$$

Then we deduce that

$$\oint_{\pi^B} |F_\pi(\boldsymbol{\alpha}; d)|^{2s}\, \mathrm{d}\boldsymbol{\alpha} \ll \langle \pi \rangle^{sc+B\varepsilon} \rho_0(\mathfrak{a})^{-2} \sum_{\xi \pmod{\pi^c}} \rho_c(\xi)^2 \ll \langle \pi \rangle^{(2s\tau+\varepsilon)B}.$$

Recall that τ was chosen sufficiently small in terms of s and k. Then it follows that for each positive number δ, one has

$$\oint |F_\pi(\boldsymbol{\alpha}; d)|^{2s}\, \mathrm{d}\boldsymbol{\alpha} \ll (q^d)^\delta,$$

whence, on recalling (17.2),

$$\begin{aligned} \oint |F(\boldsymbol{\alpha}; d)|^{2s}\, \mathrm{d}\boldsymbol{\alpha} &\ll 1 + \oint |F_0(\boldsymbol{\alpha}; d)|^{2s}\, \mathrm{d}\boldsymbol{\alpha} \\ &\ll 1 + (q^d)^\delta \sum_{\pi \in \mathcal{P}} 1 \ll (q^d)^{2\delta}. \end{aligned}$$

The conclusion of the theorem follows on recalling the definitions of $F(\boldsymbol{\alpha}; d)$ and the weight ρ_0. □

References

[1] T. C. Anderson, B. Cook, K. Hughes and A. Kumchev, *Improved ℓ^p-boundedness for integral k-spherical maximal functions*, Discrete Analysis, May 29, 2018; available as arXiv:1707.08667.

[2] G. I. Arkhipov, V. N. Chubarikov and A. A. Karatsuba, *Trigonometric sums in number theory and analysis*, de Gruyter Expositions in Mathematics, **39**, Walter de Gruyter, Berlin, 2004.

[3] R. C. Baker, *Diophantine inequalities*, London Mathematical Society Monographs, New Series, **1**, Oxford University Press, New York, 1986.

[4] B. J. Birch, *Waring's problem in algebraic number fields*, Proc. Cambridge Philos. Soc. **57** (1961), 449–459.

[5] V. Blomer and J. Brüdern, *The number of integer points on Vinogradov's quadric*, Monatsh. Math. **160** (2010), no. 3, 243–256.

[6] J. Bourgain, *On the Vinogradov mean value*, Tr. Mat. Inst. Steklova **296** (2017), Analiticheskaya i Kombinatornaya Teoriya Chisel, 36–46.

[7] J. Bourgain, C. Demeter and S. Guo, *Sharp bounds for the cubic Parsell-Vinogradov system in two dimensions*, Adv. Math. **320** (2017), 827–875.

[8] J. Bourgain, C. Demeter and L. Guth, *Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three*, Ann. of Math. (2) **184** (2016), no. 2, 633–682.
[9] D. Brink, H. Godinho and P. H. A. Rodrigues, *Simultaneous diagonal equations over* $\mathfrak{p}$*-adic fields*, Acta Arith. **132** (2008), no. 4, 393–399.
[10] J. Brüdern, *A problem in additive number theory*, Math. Proc. Cambridge Philos. Soc. **103** (1988), no. 1, 27–33.
[11] J. Brüdern and O. Robert, *Rational points on linear slices of diagonal hypersurfaces*, Nagoya Math. J. **218** (2015), 51–100.
[12] E. Croot and D. Hart, *h-fold sums from a set with few products*, SIAM J. Discrete Math. **24** (2010), no. 2, 505–519.
[13] Y. Eda, *On the mean-value theorem in an algebraic number field*, Japan. J. Math. **36** (1967), 5–21.
[14] Y. Eda, *On Waring's problem in algebraic number fields*, Rev. Colombiana Mat. **9** (1975), no. 2, 29–73.
[15] K. B. Ford, *Vinogradov's integral and bounds for the Riemann zeta function*, Proc. London Math. Soc. (3) **85** (2002), no. 3, 565–633.
[16] K. B. Ford and T. D. Wooley, *On Vinogradov's mean value theorem: strongly diagonal behaviour via efficient congruencing*, Acta Math. **213** (2014), no. 2, 199–236.
[17] S. Guo, *On a binary system of Prendiville: The cubic case*, preprint available as arXiv:1701.06732.
[18] D. R. Heath-Brown, *The cubic case of Vinogradov's mean value theorem — a simplified approach to Wooley's "efficient congruencing"*, available as arXiv:1512.03272.
[19] L.-K. Hua, *On Waring's problem*, Quart. J. Math. Oxford **9** (1938), 199–202.
[20] L.-K. Hua, *On Tarry's problem*, Quart. J. Math. Oxford **9** (1938), 315–320.
[21] L.-K. Hua, *Improvement of a result of Wright*, J. London Math. Soc. **24** (1949), 157–159.
[22] L.-K. Hua, *An improvement of Vinogradovs mean-value theorem and several applications*, Quart. J. Math. Oxford **20** (1949), 48–61.
[23] L.-K. Hua, *Additive theory of prime numbers*, American Math. Soc., Providence, RI, 1965.
[24] O. Körner, *Über Mittelwerte trigonometrischer Summen und ihre Anwendung in algebraischen Zahlkörpern*, Math. Ann. **147** (1962), 205–239.
[25] I. M. Kozlov, *The mean value theorem of I. M. Vinogradov for Gaussian numbers*, Proceedings of the IV International Conference "Modern Problems of Number Theory and its Applications" (Tula, 2001), Chebyshevskiĭ Sb. **1** (2001), 25–39.
[26] R. M. Kubota, *Waring's problem for* $\mathbb{F}_q[x]$, Dissertationes Math. (Rozprawy Mat.) **117** (1974), 60pp.
[27] W. Kuo, Y.-R. Liu and X. Zhao, *Multidimensional Vinogradov-type estimates in function fields*, Canad. J. Math. **66** (2014), no. 4, 844–873.
[28] E. Landau, *Neuer Beweis des Primzahlsatzes und Beweis des Primidealsatzes*, Math. Ann. **56** (1903), no. 4, 645–670.
[29] S. T. Parsell, S. M. Prendiville and T. D. Wooley, *Near-optimal mean value estimates for multidimensional Weyl sums*, Geom. Funct. Anal. **23** (2013), no. 6, 1962–2024.
[30] S. T. Parsell and T. D. Wooley, *A quasi-paucity problem*, Michigan Math. J. **50** (2002), no. 3, 461–469.
[31] L. G. Peck, *Diophantine equations in algebraic number fields*, Amer. J. Math. **71** (1949), 387–402.
[32] L. Pierce, *The Vinogradov mean value theorem (after Wooley, and Bourgain, Demeter and Guth)*, Séminaire Bourbaki, 69ême année, 2016–2017, pp 1134–1179, Juin 2017.
[33] Rogovskaya, N. N., *An asymptotic formula for the number of solutions of a system of equations*, Diophantine Approximations, Part II, Moskov. Gos. Univ., Moscow, 1986, pp. 78–84.

[34] P. Salberger and T. D. Wooley, *Rational points on complete intersections of higher degree, and mean values of Weyl sums*, J. London Math. Soc. (2) **82** (2010), no. 2, 317–342.
[35] P. N. Sorokin, *The mean value theorem of I. M. Vinogradov for a trigonometric sum in Gaussian numbers*, Vestnik Moscov. Univ. Ser. I Mat. Mekh. (2007), no. 6, 63–65.
[36] R. S. Steiner, *Effective Vinogradov's mean value theorem via efficient boxing*, preprint available as arXiv:1603.02536v2.
[37] E. C. Titchmarsh, *The theory of the Riemann zeta-function*, Second edition, Edited and with a preface by D. R. Heath-Brown, The Clarendon Press, Oxford University Press, New York, 1986.
[38] R. C. Vaughan, *The Hardy-Littlewood method*, Cambridge University Press, Cambridge, 1997.
[39] R. C. Vaughan and T. D. Wooley, *A special case of Vinogradov's mean value theorem*, Acta Arith. **79** (1997), no. 3, 193–204.
[40] I. M. Vinogradov, *New estimates for Weyl sums*, Dokl. Akad. Nauk SSSR **8** (1935), 195–198.
[41] I. M. Vinogradov, *The method of trigonometrical sums in the theory of numbers*, Trav. Inst. Math. Stekloff **23** (1947), 109pp.
[42] Y. Wang, *Diophantine equations and inequalities in algebraic number fields*, Springer-Verlag, Berlin, 1991.
[43] T. D. Wooley, *On Vinogradov's mean value theorem*, Mathematika **39** (1992), no. 2, 379–399.
[44] T. D. Wooley, *A note on symmetric diagonal equations*, Number theory with an emphasis on the Markoff spectrum (Provo, UT, 1991), Editors: A. D. Pollington and W. Moran, Dekker, New York, 1993, pp. 317–321.
[45] T. D. Wooley, *A note on simultaneous congruences*, J. Number Theory **58** (1996), no. 2, 288–297.
[46] T. D. Wooley, *On the local solubility of Diophantine systems*, Compositio Math. **111** (1998), no. 2, 149–165.
[47] T. D. Wooley, *Vinogradov's mean value theorem via efficient congruencing*, Ann. of Math. (2) **175** (2012), no. 3, 1575–1627.
[48] T. D. Wooley, *The asymptotic formula in Waring's problem*, Internat. Math. Res. Notices **2012** (2012), no. 7, 1485–1504.
[49] T. D. Wooley, *Vinogradov's mean value theorem via efficient congruencing, II*, Duke Math. J. **162** (2013), no. 4, 673–730.
[50] T. D. Wooley, *Translation invariance, exponential sums, and Waring's problem*, Proceedings of the International Congress of Mathematicians, August 13–21, 2014, Seoul, Korea, Volume II, Kyung Moon Sa Co. Ltd., Seoul, Korea, 2014, pp. 505–529.
[51] T. D. Wooley, *Mean value estimates for odd cubic Weyl sums*, Bull. Lond. Math. Soc. **47** (2015), no. 6, 946–957.
[52] T. D. Wooley, *Rational solutions of pairs of diagonal equations, one cubic and one quadratic*, Proc. London Math. Soc. (3) **110** (2015), no. 2, 325–356.
[53] T. D. Wooley, *Corrigendum: "The asymptotic formula in Waring's problem"*, Internat. Math. Res. Notices **2015** (2015), no. 20, 10702.
[54] T. D. Wooley, *Multigrade efficient congruencing and Vinogradov's mean value theorem*, Proc. London Math. Soc. (3) **111** (2015), no. 3, 519–560.
[55] T. D. Wooley, *The cubic case of the main conjecture in Vinogradov's mean value theorem*, Adv. Math. **294** (2016), 532–561.
[56] T. D. Wooley, *Perturbations of Weyl sums*, Internat. Math. Res. Notices **2016** (2016), no. 9, 2632–2646.
[57] T. D. Wooley, *Solvable points on smooth projective varieties*, Monatsh. Math. **180** (2016), no. 2, 391–403.

[58] T. D. Wooley, *Approximating the main conjecture in Vinogradov's mean value theorem*, Mathematika **63** (2017), no. 1, 292–350.
[59] T. D. Wooley, *Discrete Fourier restriction via efficient congruencing*, Internat. Math. Res. Notices **2017** (2017), no. 5, 1342–1389.
[60] E. M. Wright, *The Prouhet-Lehmer problem*, J. London Math. Soc. **23** (1948), 279–285.

School of Mathematics, University of Bristol, University Walk, Clifton, Bristol BS8 1TW, United Kingdom

E-mail address: matdw@bristol.ac.uk